

UNIVERSITÉ DE SHERBROOKE

Les conditions favorisant la mise en œuvre d'une évaluation en situation authentique
en sciences humaines au collégial

par
Johanne Paquin

Essai présenté à la Faculté d'éducation
en vue de l'obtention du grade de
Maître en enseignement (M.Éd.)
Maîtrise en enseignement au collégial

Septembre 2018
©Paquin Johanne, 2018

UNIVERSITÉ DE SHERBROOKE

Faculté d'éducation

Les conditions favorisant la mise en œuvre d'une évaluation en situation authentique
en sciences humaines au collégial

par

Johanne Paquin

a été évalué par un jury composé des personnes suivantes :

Julie Lyne Leroux

Julie Lyne Leroux

Denyse Lemay

Évaluatrice de l'essai

Essai accepté le 20 novembre 2018

SOMMAIRE

La mise en place du Renouveau de l'enseignement collégial en 1993 a contribué à transformer les pratiques enseignantes, mais son plein potentiel ne semble pas avoir été atteint dans tous les domaines de la relation pédagogique. Au collégial, plusieurs autrices et auteurs soulignent que les pratiques d'évaluation traditionnelles persistent. Qu'est-ce qui explique cette situation ?

Cet essai cible spécifiquement les enjeux entourant la mise en place d'une épreuve terminale de cours en situation authentique (ÉTCSA) dans le cadre des cours de première année du DEC en sciences humaines, et tente de répondre à la question suivante : quelles sont les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'évaluation en situation authentique dans des cours de première année du programme préuniversitaire en sciences humaines au collégial ? Pour y arriver, l'objectif de recherche suivant a été retenu : documenter les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans des cours de première année du programme d'études préuniversitaires en sciences humaines au collégial.

Pour ce faire, une étude de cas unique a été réalisée auprès d'un enseignant du programme de sciences humaines ayant expérimenté la conception et la mise en œuvre d'une ÉTCSA dans le cadre d'un cours de première année de ce programme. Il a complété un questionnaire, participé à une entrevue semi-dirigée et fourni un dossier d'artefact, qui ont constitué les données analysées à partir d'une approche qualitative.

Les résultats de la collecte de données ont été présentés et analysés par phase de mise en œuvre d'une ÉTCSA, soit l'élaboration, l'administration et l'évaluation.

Pour la phase d'élaboration, quatre enjeux majeurs ont été identifiés : le temps, l'interaction avec des collègues, la disponibilité de ressources spécialisées et la latitude sur le plan institutionnel. La phase d'élaboration est cruciale et doit être réalisée minutieusement en raison des impacts sur les deux phases subséquentes.

Pour la phase d'administration, les enjeux sont différents lors de la première mise en œuvre et lors des mises en œuvre subséquentes. La première mise en œuvre est particulièrement exigeante et présente les enjeux spécifiques suivants : les difficultés et besoins non anticipés des étudiantes et étudiants et la maîtrise imparfaite de la démarche et du matériel de la part de l'enseignante ou de l'enseignant. Lors des mises en œuvre subséquentes, certains enjeux récurrents ou contextuels ont été ciblés, dont un besoin accru d'accompagnement des étudiantes et étudiants, l'impact de la variabilité du contexte pédagogique et des imprévus, et l'importance des ressources humaines et matérielles suffisantes.

Pour la phase d'évaluation, trois enjeux sont identifiés. Sur le plan de l'équité, l'hétérogénéité des productions résultant de la démarche et l'impact des comportements et des choix des étudiantes et étudiants sur les autres étudiantes et étudiants durant le déroulement de l'ÉTCSA ont été soulevés. Il y a également un écart entre la première mise en œuvre, plus complexe du fait des imprévus et des ajustements en découlant, et les subséquentes. Concernant les enjeux en temps et en ressources, outre la première mise en œuvre, adopter l'ÉTCSA n'est pas plus exigeante pour la phase d'évaluation qu'une épreuve terminale de cours traditionnelle.

Finalement, le participant à cette étude a rencontré plusieurs des enjeux recensés dans la littérature à propos de la mise en œuvre d'une ÉTCSA, pour les quatre catégories présentées dans le cadre de référence, soit la qualité et rigueur du processus évaluatif, les exigences en temps et en ressource, le contexte et les contraintes institutionnelles et les problèmes de nature pédagogique.

TABLE DES MATIÈRES

SOMMAIRE	4
LISTE DES TABLEAUX	10
REMERCIEMENTS	11
INTRODUCTION.....	12
PREMIER CHAPITRE - LA PROBLÉMATIQUE.....	14
1. LE CONTEXTE DE LA RECHERCHE	14
1.1 Les pratiques évaluatives au collégial depuis le Renouveau pédagogique	14
1.2. La lente transformation des pratiques évaluatives	16
1.3. Un élément ponctuel : le renouvellement du corps professoral	19
2. LE PROBLÈME DE RECHERCHE	19
3. LA QUESTION GÉNÉRALE	23
DEUXIÈME CHAPITRE - LE CADRE DE RÉFÉRENCE.....	26
1. ORIGINE ET GÉNÉALOGIE DU CONCEPT D'ÉVALUATION EN SITUATION AUTHENTIQUE	26
1.1. La remise en question de l'approche psychométrique en évaluation et l'émergence du courant de l'évaluation des performances	27
1.2. De l'évaluation des performances à l'évaluation en situation authentique	29
2. LE CONCEPT D'ÉSA DANS LA LITTÉRATURE.....	31
2.1. Problèmes sémantiques et de définition.....	31
2.2. Spécificités des corpus anglophone et francophone	33
3. QU'EST-CE QU'UNE ÉVALUATION EN SITUATION AUTHENTIQUE ?	36
3.1. Le modèle de Newmann et Archbald	36

3.1.1.	Un modèle tridimensionnel.....	37
3.1.2.	Le contexte de réalisation d'une ÉSA selon Archbald et Newmann.....	43
3.1.3.	Un 5 ^e critère contextuel : le contexte de réalisation permet l'évolution de l'apprentissage.....	44
3.1.4.	L'évaluation d'une ÉSA selon le modèle d'Archbald et Newmann.....	47
3.2.	Le modèle d'ÉSA selon Wiggins :	48
3.2.1.	La définition de l'ÉSA et ses caractéristiques fondamentales selon Wiggins	49
3.2.2.	Un modèle quadridimensionnel à l'usage des praticiennes et praticiens.....	51
3.3.	Comparaison des modèles d'Archbald et Newmann et de Wiggins : deux constats pertinents pour notre démarche de recherche	55
4.	LES DIFFICULTÉS RENCONTRÉES DANS LA MISE EN ŒUVRE DE L'ÉSA : QU'EN DIT LA LITTÉRATURE ?	58
4.1.	La qualité et la rigueur du processus évaluatif :	58
4.2.	Les exigences en temps et en ressources	61
4.3.	Le contexte et les contraintes institutionnelles :	67
4.4.	Les problèmes de nature pédagogique :	68
5.	LES OBJECTIFS SPÉCIFIQUES DE L'ESSAI	69
	TROISIÈME CHAPITRE - LA MÉTHODOLOGIE	73
1.	TYPE D'ESSAI.....	73
2.	LA POSTURE ÉPISTÉMOLOGIQUE.....	76
2.1	La vision de la réalité	77
2.2	La nature du savoir et la place du chercheur	77
2.3	Finalité de la recherche.....	79
3.	TYPE DE RECHERCHE ET APPROCHE MÉTHODOLOGIQUE	80
3.1	Les critères de sélection.....	81
3.2	Le participant.....	84
3.3	Méthode de collecte des données	86

3.3.1. L'entrevue semi-dirigée	86
3.3.2. Le questionnaire	89
3.3.3. Le dossier d'artefacts	89
3.4. La méthode d'analyse et de traitement des données	90
4. LES ASPECTS ÉTHIQUES	93
4.1. Respect de la dignité humaine	93
4.2. Respect du consentement libre et éclairé	93
4.3. Respect des personnes vulnérables	94
4.4. Respect de la vie privée et des renseignements personnels	94
5. LES MOYENS PRIS POUR ASSURER LA SCIENTIFICITÉ	96
5.1. La crédibilité (aussi appelée validité interne):	96
5.2. La transférabilité (aussi appelé validité externe)	98
5.3. La fiabilité	99
5.4. La confirmabilité	100
QUATRIÈME CHAPITRE - LA PRÉSENTATION ET L'INTERPRÉTATION DES RÉSULTATS	101
1. PRÉSENTATION ET ANALYSE DES RÉSULTATS	102
1.1. Phase d'élaboration	103
1.1.1. Le besoin de temps	104
1.1.2. L'interaction avec des collègues	105
1.1.3. Disponibilité de ressources spécialisées	107
1.1.4. Le contexte institutionnel et académique	108
1.2. Phase d'administration	110
1.2.1. Les enjeux associés à la première mise en œuvre	112
1.2.2. Les enjeux récurrents ou contextuels	115
1.3. Phase d'évaluation	120
1.3.1. Les enjeux d'équité	121
1.3.2. Enjeux liés à la non-standardisation des tâches réalisées par les étudiantes et étudiants	121
1.3.3. Impacts des comportements et des choix des étudiantes et étudiants sur l'équité	122

1.3.4.	Les enjeux concernant le bagage socioculturel	124
1.3.5.	Les enjeux liés à la première mise en œuvre	125
1.3.6.	Les enjeux liés au temps et aux ressources.....	126
2.	ANALYSE ET INTERPRÉTATION DES RÉSULTATS	127
2.1.	La qualité et la rigueur du processus évaluatif	128
2.2.	Exigences en temps et en ressources	131
2.3.	Le contexte et les contraintes institutionnelles.....	134
2.4.	Les problèmes de nature pédagogique	136
	CONCLUSION.....	139
1.	LES ÉLÉMENTS CENTRAUX DE CET ESSAI	139
2.	LES LIMITES DE LA DÉMARCHE.....	149
3.	LES PISTES POUR DES DÉMARCHES FUTURES	148
	RÉFÉRENCES BIBLIOGRAPHIQUES.....	152
	ANNEXE A - FORMULAIRE D'INFORMATION ET DE CONSENTEMENT	161
	ANNEXE B - OUTIL DE COLLECTE : LE QUESTIONNAIRE.....	166
	ANNEXE C - OUTIL DE COLLECTE : LE SCHÉMA D'ENTREVUE.....	172
	ANNEXE D - OUTIL DE COLLECTE : GRILLE DE COLLECTE DES ARTEFACTS	181
	ANNEXE E - TABLEAU 1 – SYNTHÈSE DES PROBLÈMES RELIÉS À LA CONCEPTION ET LA MISE EN ŒUVRE D'UNE ÉSA RECENSÉS DANS LA LITTÉRATURE.....	2
	ANNEXE F - TABLEAU 2 – RÉSUMÉ DES ENJEUX, DIFFICULTÉS, CONDITIONS ET ACTIONS PERÇUES COMME FAVORABLES PAR JULIEN POUR LA PHASE D'ADMINISTRATION DE L'ÉTCSA	6
	ANNEXE G - TABLEAU 3 – RÉSUMÉ DES ENJEUX, DIFFICULTÉS, CONDITIONS ET ACTIONS PERÇUES COMME FAVORABLES PAR JULIEN POUR LA PHASE D'ADMINISTRATION DE L'ÉTCSA	10
	ANNEXE H - TABLEAU 4 – RÉSUMÉ DES ENJEUX, DIFFICULTÉS, CONDITIONS ET ACTIONS PERÇUES COMME FAVORABLES PAR JULIEN POUR LA PHASE D'ÉVALUATION DE L'ÉTCSA.....	15

LISTE DES TABLEAUX

Tableau 1: Synthèse des problèmes reliés à la conception et la mise en œuvre d'une ÉSA.....	2
Tableau 2: Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase de conception de l'ÉTCSA.....	6
Tableau 3: Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'administration de l'ÉTCSA	10
Tableau 4: Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'évaluation de l'ÉTCSA.....	15

REMERCIEMENTS

Cet essai n'aurait pu être possible sans les précieux conseils et le soutien de ma directrice, Julie Lyne Leroux, qui a su me guider sur le plan théorique et conceptuel et maintenir ma motivation tout au long du processus, et le Programme de formation pour l'obtention d'une maîtrise FNEEQ-CSN (Fédération Nationale des Enseignantes et Enseignants du Québec – Confédération des Syndicats Nationaux), qui m'a permis d'avoir le temps nécessaire à la rédaction du cadre théorique et de la méthodologie de cette recherche.

Et surtout, je souhaite exprimer toute ma reconnaissance envers mes parents, Pauline Désaulniers et Réjean Paquin, qui m'ont transmis des valeurs essentielles à la réussite d'études universitaires – la persévérance, l'autonomie et le goût du travail bien fait - et m'ont inébranlablement encouragé et soutenu dans mon parcours académique et professionnel : c'est le plus précieux des héritages, et je leur en serai toujours reconnaissante.

INTRODUCTION

Depuis la mise en place du Renouveau de l'enseignement collégial en 1993, une petite révolution pédagogique a eu lieu dans les cégeps, et les sciences de l'éducation y ont grandement contribué. Si cette réforme, inspirée des paradigmes constructiviste et socioconstructiviste, a contribué à transformer les pratiques enseignantes au quotidien, son plein potentiel ne semble pas avoir été atteint dans tous les domaines de la relation pédagogique. C'est notamment le cas des pratiques évaluatives, qui doivent encore être améliorées dans de nombreuses institutions collégiales et pour lesquelles les transformations des pratiques se font plus lentement (Gouvernement du Québec, 2009, 2012; Bélanger et Tremblay, 2012).

Afin de mieux comprendre cette situation, cet essai s'intéresse aux facteurs qui favorisent et défavorisent la mise en œuvre d'évaluation terminale de cours en situation authentique, une approche compatible avec les orientations de la réforme et de l'approche par compétence, en posant la question suivante : quelles sont les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'évaluation en situation authentique dans des cours de première année du programme préuniversitaire en sciences humaines au collégial ?

Pour y répondre, l'objectif de recherche principal suivant a été retenu : documenter les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans des cours de première année du programme d'études préuniversitaires en sciences humaines au collégial.

Le premier chapitre présente la problématique associée à cette question et cet objectif : on y retrouve la présentation du contexte et des enjeux associés, puis du

problème de recherche, suivi de la question de recherche, de l'objectif principal de recherche et des trois objectifs secondaires de recherche retenus.

Le second chapitre présente le cadre de référence sur lequel s'appuie la démarche. L'origine et la généalogie du concept d'Évaluation en situation authentique (ÉSA) et sa place dans la littérature y sont abordés, suivi d'une présentation détaillée des deux principaux modèles d'ÉSA utilisés, celui d'Archbald et Newmann et celui de Wiggins. Finalement, une synthèse des difficultés rencontrées lors de la mise en œuvre d'ÉSA tirées de la littérature est présentée.

Le troisième chapitre présente la méthodologie utilisée pour la collecte de données. La position épistémologique privilégiée, le type de méthodologie, l'ensemble de la démarche concernant le cas d'étude choisi et les précautions éthiques adoptées y sont détaillés.

Le quatrième chapitre présente les résultats de la collecte de données. Pour chacune des étapes de la mise en œuvre d'une ÉSA – élaboration, administration et évaluation – les données sont présentées et analysées. Ensuite, elles sont interprétées en regard des difficultés associées à la mise en œuvre d'une ÉSA telles que recensées dans la littérature, et présentées dans le cadre de référence.

En conclusion, une synthèse des éléments et résultats principaux de cette recherche est présentée, suivi d'une réflexion sur les limites de cet essai et des propositions et pistes pour de futures recherches en lien avec la réalisation d'épreuves terminales de cours en situation authentique.

PREMIER CHAPITRE

LA PROBLÉMATIQUE

Afin de bien situer cet essai dans le contexte plus global des savoirs liés aux sciences de l'éducation, ce chapitre présente d'abord le contexte et les enjeux ayant mené à la réalisation de cet essai. S'en suit la présentation du problème de recherche, ainsi que la question et l'objectif de recherche. Finalement, il sera question du positionnement de cette démarche parmi les pôles de recherche de la Maîtrise en enseignement.

1. LE CONTEXTE DE LA RECHERCHE

Depuis la création des cégeps, les pratiques évaluatives ont grandement évolué, tout comme la composition du corps professoral. Il importe donc de présenter les éléments centraux de cette évolution.

1.1 Les pratiques évaluatives au collégial depuis le Renouveau pédagogique

En 1993, à la suite des travaux de la Commission parlementaire de l'éducation réalisés l'année précédente, le Ministère de l'Enseignement Supérieur et des Sciences (MESS) a demandé aux cégeps d'instaurer une réforme majeure intitulée *Renouveau de l'enseignement collégial*. En plus d'importants impacts sur les modes de gestion et sur la reddition de compte des établissements, cette réforme commandait une transformation des programmes et de l'enseignement au niveau collégial (Boisvert, Lacoursière et Lallier, 2006). Au cœur de nombreux changements radicaux quant à la manière de planifier les programmes de formation au collégial se trouve le concept d'«approche par compétences», un «mode d'élaboration de cours ou de programmes

d'études qui consiste à définir les compétences inhérentes à l'exercice d'une profession et à les formuler, dans les programmes, en objectifs et standards» (Legendre, 2005).

Parmi les enseignantes et les enseignants du collégial, les réactions à ces nouvelles conceptions pédagogiques furent multiples. Outre les réserves quant aux fondements épistémologiques exprimés par certains (Gruda, 1994; Ouimet, 1994*c*), le manque d'assises théoriques claires sur lesquels pouvaient se baser les enseignantes et les enseignants (Goulet, 1994; Gouvernement du Québec, 2009, 2012; Langevin, Boily et Talbot, 2004) (Carpentier et Conseil supérieur de l'éducation du Québec, 2000), les réticences dues au manque de consultation préimplantation (Goulet, 1994; Gouvernement du Québec, 2009, 2012; Langevin *et al.*, 2004) et une impression d'improvisation due aux courts délais (Goulet, 1994; Gouvernement du Québec, 2004; Ouimet, 1994*a*, 1994*b*, 1994*d*), les enseignantes et les enseignants du collégial faisaient face à un défi de taille en raison des nécessaires et importantes modifications dans la pratique enseignante qu'impliquait cette réforme. Si les programmes techniques ont adopté plus rapidement et facilement cette nouvelle approche pédagogique, la transition fut plus longue et difficile du côté de la formation générale et des programmes préuniversitaires (Boutin, 2006; Langevin *et al.*, 2004).

Vingt ans plus tard, il semble que ce défi a été grandement relevé. Alors qu'en 1993, Howe et Ménard dénonçaient le manque de connaissance des paradigmes constructiviste et socioconstructiviste chez les enseignantes et les enseignants et soulignaient leurs lacunes importantes notamment en matière d'évaluation, les recherches plus récentes démontrent que les pratiques pédagogiques ont sensiblement évolué depuis et que les orientations du Renouveau pédagogique et de l'approche par compétence ont lentement, mais durablement, transformé les pratiques pédagogiques des enseignantes et des enseignants du collégial (Bélanger et Tremblay, 2012; Chbat, 2004, 2005, 2006; Gouvernement du Québec, 2009; Isabel, 2000; Leroux, 2010; Pineault, 2001). La Commission d'évaluation de l'enseignement collégial (CÉEC) confirmait d'ailleurs ce constat dans son bilan réalisé en 2009 : les principes du

Renouveau pédagogique ont largement été appliqués dans les institutions et bien que quelques ajustements restent à faire, les effets se font concrètement sentir sur le plan des approches pédagogiques adoptées par les enseignantes et les enseignants. Un domaine de l'enseignement fait toutefois exception : l'évaluation. La CÉEC (Gouvernement du Québec, 2009, 2012) a documenté une certaine continuité des pratiques d'évaluation non conforme aux principes du Renouveau pédagogique, et la recherche de Bélanger et Tremblay (2012) témoigne de la persistance de certaines croyances difficilement compatibles avec l'approche par compétence, malgré le constat d'une transformation de certaines pratiques.

1.2. La lente transformation des pratiques évaluatives

Dès le début du Renouveau pédagogique, l'évaluation ressort comme un aspect particulièrement sensible et son adaptation à l'approche par compétence ne va pas de soi pour les Collèges et les enseignantes et les enseignants. En 1995, en amont du Renouveau au collégial, la CÉEC soulevait que dans plusieurs institutions, certaines règles institutionnelles en matière d'évaluation étaient même incompatibles avec les principes de l'approche par compétence :

L'adoption de l'approche par compétences exige une réflexion sur la façon d'évaluer la maîtrise des compétences et, dans bien des cas, elle entraîne des modifications substantielles aux pratiques existantes. L'état actuel des PIEA de plusieurs collèges montre que l'implantation de cette approche ne se fait pas sans difficulté. De fait, l'inadaptation des règles d'évaluation des apprentissages à ce nouveau contexte a représenté le problème le plus fréquent amenant des recommandations de la part de la Commission. Deux raisons expliquent les difficultés rencontrées par les collèges. 1 – Selon le ministère de l'Éducation, l'évaluation de la compétence doit se faire dans un contexte de réalisation précis qu'il définit comme *«les conditions dans lesquelles sera placé l'élève au moment de démontrer, a posteriori, l'atteinte d'une compétence»*. C'est dire que l'élève ne devrait normalement pas pouvoir faire cette démonstration avant d'avoir complété une portion très substantielle du cours. Or, cette façon d'évaluer la maîtrise des compétences n'est pas facile à réconcilier avec la tradition d'évaluation continue qui s'est développée dans le réseau collégial. En effet, de nombreux collèges se sont donné des règles favorisant le recours à

plusieurs évaluations (tests, travaux, examens...) au cours d'une session plutôt qu'à une seule, et limitant le poids de chacune des évaluations sommatives. Certaines règles par exemple stipulent qu'aucune évaluation ne peut compter pour plus de 30 % ou 40 % ou 50 % de la note finale. De telles règles visent habituellement des objectifs fort louables comme amener l'élève à travailler régulièrement au cours de la session ou lui permettre de se situer par rapport à ses progrès. Mais elles peuvent avoir pour effet qu'un élève obtienne la note de passage à un cours sans avoir démontré qu'il maîtrise réellement la compétence visée ou vice versa. (Gouvernement du Québec, 1995)

On constate donc à quel point les transformations en matière d'évaluation sont complexes et lentes et qu'elle constitue un point de tension dès le départ : à preuve, deux ans après l'implantation de la réforme, les politiques formelles ou en d'autres termes la Politique institutionnelle d'évaluation des apprentissages (PIEA) de chaque cégep exigeait toujours des pratiques évaluatives incohérentes avec les orientations officiellement prônées par le ministère de l'Éducation. (Gouvernement du Québec, 1995)

À la suite de ces constats, les institutions collégiales ont modifié leur cadre réglementaire. En 2004, la CÉEC publiait un nouveau rapport attestant que les cégeps ont corrigé ces problèmes dans leur PIEA, mais qu'il semble que les effets de ces modifications n'aient pas totalement percolé jusque dans les classes sur le plan des pratiques évaluatives (Gouvernement du Québec, 2004).

Ainsi, bien que le renouveau pédagogique ait eu des effets substantiels sur le plan des pratiques d'enseignement et de l'harmonisation des politiques des collèges, sur le terrain, dans les salles de classe, la pratique n'a pas complètement suivi. La CÉEC (2012) lance donc une démarche d'évaluation institutionnelle centrée sur l'application des PIEA dans les collèges du Québec, démarche au cours de laquelle chaque institution doit « s'assurer que sa politique est appliquée conformément à ce que le texte prévoit, notamment en ce qui concerne l'exercice des responsabilités, et qu'elle permet l'atteinte des objectifs poursuivis » (p. 9) et au terme de laquelle est publié le

rapport synthèse *L'application des Politiques institutionnelles d'évaluation des apprentissages : Pour des évaluations justes et équitables* (2012). Près de 20 ans après la réforme, la CEEC (2012) constate que dans l'ensemble, les collèges ont de bonnes pratiques en matière d'évaluation. Elle fait toutefois le constat suivant :

La Commission a toutefois relevé des lacunes importantes au regard de l'évaluation sommative et elles concernent un bon nombre de collèges. L'examen fait par la Commission a souligné que certaines évaluations finales de cours ne sont pas de nature synthèse et ne portent pas sur la globalité de la compétence. Le niveau taxonomique de l'épreuve n'est pas toujours approprié aux standards visés. La Commission a également constaté que certaines épreuves consistent en un travail d'équipe qui ne comporte aucune mesure individuelle permettant à l'étudiant de démontrer qu'il a atteint l'objectif du cours. Enfin, elle a noté des cas où la pondération de l'épreuve terminale n'est pas significative, de sorte que sa réussite ne permet pas d'attester l'atteinte des objectifs du cours selon les standards et ainsi la maîtrise de la compétence. La Commission a donc formulé plusieurs avis aux collèges afin que l'ensemble des épreuves finales de cours permette de mesurer, pour chacun des étudiants, le niveau d'atteinte des objectifs selon les standards visés, autant à la formation ordinaire qu'à la formation continue. La Commission constate que la mesure du niveau d'atteinte des objectifs d'un cours selon les standards est une forme d'évaluation dont l'implantation doit être poursuivie. (CÉEC, 2012, p.16)

Ainsi, il y a une lente, mais réelle évolution des pratiques évaluatives des enseignantes et des enseignants, notamment par l'introduction d'évaluations formatives, rarement pratiquées avant la réforme (Howe et Ménard, 1993). Mais un dernier bastion demeure : l'évaluation terminale des cours, incarnée dans le contexte collégial québécois par l'épreuve terminale de cours. À la suite de ce rapport, la CÉEC interpelle les collèges spécifiquement sur cette question afin que chaque cours intègre une épreuve synthèse conforme aux principes de l'approche par compétence et demande que des modifications soient apportées par les enseignantes et les enseignants afin que la démarche évaluative de chaque cours certifie réellement les capacités des étudiantes et des étudiants à intégrer les apprentissages dans des tâches complexes, authentiques et significatives (CÉEC, 2012).

1.3. Un élément ponctuel : le renouvellement du corps professoral

Finalement, mentionnons qu'un élément ponctuel s'ajoute au contexte dans lequel se situe le présent essai, soit l'importante transformation de la composition du corps enseignant. Dès 2008, la Fédération des cégeps attirait l'attention de la communauté collégiale sur l'importance de ce phénomène à venir, qui est maintenant en cours : elle soulignait qu'en 2006-2007, 37,1 % du corps enseignant avaient 50 ans ou plus (Fédération des cégeps et Institut du Nouveau Monde, 2008, pp. 13-14), d'où l'importance de se préparer aux bouleversements engendrés par un renouvellement rapide du corps professoral.

Sur le plan de l'analyse des pratiques pédagogiques et évaluatives, cette situation ne doit pas être négligée, car certaines études dont celle de Langevin *et al.* (2004), indiquaient qu'il y a davantage de réfractaires irréductibles à l'approche par compétence chez les enseignantes et enseignants ayant débuté leur carrière avant l'introduction du Renouveau pédagogique. Quel effet aura le départ de cette génération d'enseignantes et d'enseignants sur les pratiques des nouveaux arrivants dans le métier ? Ainsi, cet élément contextuel n'est pas à négliger dans l'étude des facteurs favorisant ou non une transformation des pratiques évaluatives des enseignantes et des enseignants.

2. LE PROBLÈME DE RECHERCHE

Nous sommes donc devant un paradoxe intrigant : la littérature nous informe que dans les deux dernières décennies, les enseignantes et les enseignants ont significativement transformé leurs pratiques d'enseignement et leur approche didactique en s'insérant dans l'approche par compétence, mais que les pratiques d'évaluation traditionnelles persistent, tout particulièrement lorsqu'il est question d'évaluation sommative et d'épreuve synthèse de programme ou d'épreuve terminale de cours. Qu'est-ce qui explique ces pratiques contradictoires chez les enseignantes et enseignants ?

Si les démarches de la CÉEC évaluent en profondeur la conformité des politiques et des pratiques aux approches pédagogiques prescrites par le Renouveau pédagogique, elle n'aborde pas les raisons de cette lente et faible intégration de ces orientations dans les classes. Or, il existe peu de littérature récente sur le sujet. L'étude de Howe et Ménard traite de cette question, mais elle date de 1993 et documente la situation au moment de l'implantation de la réforme : on ne peut donc s'y fier pour connaître la situation actuelle. Plus récemment, certaines études se sont intéressées aux transformations des pratiques pédagogiques dans leur ensemble à la suite du Renouveau pédagogique (Chbat, 2004, 2005, 2006; Langevin *et al.*, 2004), mais la question de l'évaluation n'est qu'une dimension parmi d'autres. Certaines études se sont intéressées spécifiquement aux pratiques évaluatives : c'est le cas de Leroux, qui documente de nouvelles pratiques en matière d'évaluation depuis la réforme (Leroux, 2010), mais ne traite pas des zones de résistance aux changements et de leurs causes. Cette littérature participe à une meilleure compréhension des pratiques pédagogiques et évaluatives post-réformes, mais ne permet pas de mieux comprendre en profondeur les causes de la lente adaptation de certains domaines de l'évaluation au nouveau paradigme amené par la réforme de 1993. Certains textes abordent indirectement cet enjeu en soumettant quelques hypothèses, mais celles-ci ne proviennent pas de travaux dont l'objectif premier était d'identifier de tels facteurs explicatifs. Parmi ces hypothèses sur les difficultés d'adaptation des pratiques évaluatives, on retrouve notamment celles des réticences épistémologiques et du manque de formation.

Une première hypothèse sur les raisons du repli sur les pratiques pédagogiques traditionnelles attribue cette situation au caractère réfractaire face aux approches constructivistes et socioconstructivistes et à la réforme dans son ensemble : certains enseignantes ou enseignants, notamment parmi celles et ceux enseignant dans le secteur préuniversitaire et à la formation générale et parmi les plus expérimentés (Langevin *et al.*, 2004), seraient particulièrement rébarbatifs à transformer leurs pratiques pédagogiques afin de les adapter à l'approche par compétence. Ce manque de réceptivité aurait été exacerbé par le contexte dans lequel la réforme a été implantée,

engendrant le sentiment chez les enseignantes et les enseignants que « la réforme du collégial [...] est d'abord et avant tout une réforme technocratique, une réforme décidée, pensée et organisée en dehors de ceux et de celles qui, au jour le jour, font la pédagogie collégiale et continueront de la faire: les enseignantes et les enseignants. » (Forcier, 1994, dans Tremblay, 1995, p.28), et occasionnant plusieurs frictions entre les architectes de la réforme et les enseignantes et les enseignants (Gouvernement du Québec, 2004; Laberge, 2010; Langevin *et al.*, 2004). Ce point de vue, qui s'inscrit en partie dans la continuité des travaux de Howe et Ménard (1993), mène à la conclusion que la situation relevée par la CÉEC viendrait de l'absence de volonté de la part des enseignantes et des enseignants de modifier leurs pratiques, notamment en raison de certains désaccords quant à la finalité de cette réforme et son caractère utilitariste (Boutin, 2006).

Or, sans nier que cette situation puisse exister chez certaines enseignantes ou certains enseignants, la thèse de la réticence épistémologique comme principal facteur explicatif du manque de pénétration du nouveau paradigme évaluatif est mise à mal par les constats présentés précédemment : s'il s'agissait du cœur du problème, l'ensemble des pratiques pédagogiques – et non seulement l'évaluation sommative – n'aurait pas subi de modifications. Or, les études présentées précédemment démontrent qu'il y a eu une intégration tangible des principes du Renouveau pédagogique dans les pratiques pédagogiques des enseignantes et des enseignants ainsi que dans les aspects non sommatifs de l'évaluation.

Le problème d'adaptation des évaluations sommatives est-il alors un problème de formation des enseignantes et des enseignants ? Si la formation pédagogique des enseignantes et des enseignants du collégial constitue probablement une partie de l'explication (Howe et Ménard, 1993; Laberge, 2010), la formation en pédagogie n'étant pas obligatoire, elle ne peut pas non plus être la seule. Tout d'abord, les outils théoriques et conceptuels concernant l'évaluation dans un contexte d'approche par compétence sont largement disponibles dans des ouvrages en éducation. Plusieurs

auteurs québécois et autrices québécoises ont d'ailleurs publié sur ce thème, en adaptant cette approche au contexte collégial (Durand et Chouinard, 2012; Gérard, 2008; Goulet, 1995; Goulet, 1995; Guy, Deshaies et Poirier, 2004; Leroux, 2010; Louis, 2004; Scallon, 2004) et postsecondaire (Prégent, Bernard et Kozanitis, 2009). Quelle est l'intégration effective de ces écrits dans les pratiques évaluatives des enseignantes et des enseignants? De plus, la prise de contact des enseignantes et les enseignants du collégial avec ces savoirs a été facilitée par de nombreux programmes de formation destinés aux enseignantes et aux enseignants du niveau collégial qui se sont d'ailleurs multipliés dans les dernières années : Programme court en enseignement postsecondaire de l'UQAM, DESS en enseignement collégial de l'Université Laval, Microprogramme de 2^e cycle en formation à l'enseignement postsecondaire de l'Université de Montréal et les différents programmes PERFORMA de l'Université de Sherbrooke. Certains départements et services de ressources humaines favorisent l'embauche d'enseignantes et d'enseignants détenant une telle formation, en plus d'encourager les enseignantes et enseignants déjà en poste à faire de la formation continue en pédagogie, notamment en remboursant les frais de scolarité des cours suivis, en organisant des journées pédagogiques biannuelles et en mettant à la disposition des enseignantes et des enseignants des conseillères et conseillers pédagogiques pouvant les accompagner dans leurs démarches. Conséquemment, cette situation constatée par la CÉEC ne résulte pas d'un manque de connaissances ou d'expertise, mais bien de leur intégration concrète dans les pratiques des enseignantes et des enseignants du collégial, laquelle ne s'explique pas uniquement par un manque de formation du personnel enseignant.

Ainsi, bien que ces facteurs explicatifs puissent concerner une partie des causes du problème soulevé par la CÉEC concernant les pratiques en matière d'épreuve terminale de cours, ils ne peuvent à eux seuls expliquer entièrement la situation. Qu'est-ce qui peut expliquer cette persistance des pratiques en matière d'évaluation sommative reliée au paradigme traditionnel chez des enseignantes et des enseignants qui ont toutefois adopté les orientations de la réforme pour les autres dimensions de leur

pratique pédagogique ? Et chez les enseignantes et enseignants acceptant l'esprit général du Renouveau en matière d'évaluation, qu'en est-il de leur capacité à implanter concrètement ces pratiques ? Bref, quels sont les facteurs qui favorisent ou défavorisent l'implantation de pratiques évaluatives sommatives respectant les principes de l'évaluation des compétences ? Aucune étude ne s'est penchée sur cette problématique pourtant bien réelle et ayant des impacts concrets dans le quotidien des enseignantes et des enseignants et des étudiantes et des étudiants.

3. LA QUESTION GÉNÉRALE

Il ressort donc que pour comprendre en profondeur les raisons des lacunes constatées par la CÉEC, des recherches supplémentaires sont nécessaires. En effet, si certaines pistes d'explications ont déjà été explorées, elles ne permettent pas d'expliquer entièrement le phénomène. Ainsi, il existe un besoin de documenter plus en profondeur les facteurs facilitant et entravant l'intégration de pratiques en matière d'évaluations sommatives selon les principes des compétences au collégial. Comme un tel objectif outrepasserait largement le cadre d'un essai professionnel, nous avons choisi d'aborder une dimension précise de l'évaluation sommative : l'épreuve terminale de cours en sciences humaines au collégial. Ce choix a été fait en raison des problèmes documentés en la matière et du besoin d'y remédier dans un avenir rapproché (Gouvernement du Québec, 2004; Gouvernement du Québec, 2008).

De plus, parmi les facteurs interférant dans la réalisation d'épreuves terminales de cours conformes aux principes du Renouveau pédagogique, nous documenterons plus particulièrement les conditions jugées favorables et les difficultés de la mise en œuvre d'épreuve terminale de cours en situation authentique (ÉTCSA). Le choix de cibler les conditions favorables et défavorables parmi les facteurs en cause dans la problématique des épreuves terminales de cours relève de la nécessité d'être réaliste quant à l'ampleur à donner à notre démarche de recherche dans le cadre d'un essai.

Pour ce qui est du choix de cibler l'aspect du contexte authentique parmi les critères d'une évaluation s'inscrivant dans l'approche par compétences, il repose sur le fait que c'est souvent un des aspects les plus difficiles à mettre en place dans le cadre de l'évaluation sommative. De plus, l'inclusion d'un caractère authentique dans les évaluations repose quasi entièrement entre les mains des enseignantes et enseignants en tant que concepteurs et maîtres d'œuvre des scénarios évaluatifs, et notre souhait est de concentrer notre recherche sur les aspects de l'évaluation contrôlés par les enseignantes et les enseignants.

Finalement, nous nous intéresserons à cet enjeu spécifiquement dans le cadre des cours de première année du DEC en sc. humaines. Outre le fait qu'il s'agisse du programme dans lequel nous enseignons - ce qui permettra un réinvestissement des découvertes et connaissances développées au cours de cette recherche dans notre milieu d'enseignement - ce choix découle de trois considérations établissant la pertinence de ce choix :

1. L'absence de connaissances spécifiques concernant les conditions favorisant ou défavorisant l'implantation d'ÉSA dans le domaine des sciences humaines dans le contexte collégial : en effet, outre le texte de Scheurman et Newmann (1998) qui traite de l'adaptabilité de son modèle d'ÉSA dans les cours de sciences sociales à l'université (ce qui demeure tout de même éloigné du contexte de notre étude, étant donné que les sciences sociales ne forment qu'une partie des sciences humaines, que le texte traite d'un contexte universitaire, aux États-Unis de surcroît) nous avons constaté qu'il n'existe pas de recherche ayant documenté des facteurs entravant ou facilitant l'instauration de telles évaluations dans le contexte propre au DEC en sciences humaines;
2. La place et l'importance de ce programme dans le réseau collégial québécois : en 2009, 47 307 étudiantes et étudiants étaient inscrits au DEC en sciences humaines (formation régulière), soit 26 % des effectifs collégiaux de la formation régulière (Gouvernement du Québec, 2013). Ces proportions sont relativement stables

depuis une décennie, ce qui place ce programme chaque année au premier rang en termes de nombre d'étudiantes et d'étudiants inscrits (Gouvernement du Québec, 2013). Ainsi, le choix de ce programme est tout à fait pertinent étant donné le nombre de personnes (population étudiante et personnel enseignant) touchées par cet enjeu pédagogique et des possibles retombées pédagogiques de cette recherche;

3. L'existence d'enjeux spécifiques à ce programme, particulièrement dans les cours de première année : Malgré les près de 50 000 étudiantes et étudiants inscrits chaque année au DEC en sciences humaines entre 2005 et 2009, moins de 15 000 diplômes ont été décernés annuellement durant cette période (Gouvernement du Québec, 2013). Ce programme est touché par d'importants problèmes de réussite et de persévérance scolaire, particulièrement en première année (Boisvert et Paradis, 2008; Boisvert, 2007; Fortier, 2003; Fréchette et Mignault, 2010; Lussier, 2004). On remarque aussi un taux de diplomation inférieur chez les garçons, en comparaison avec leurs collègues féminines (Boisvert et Paradis, 2008; Boisvert, 2007; Fréchette et Mignault, 2010). Parmi les facteurs identifiés, on retrouve notamment des problèmes de motivation et la difficulté à s'adapter aux méthodes pédagogiques utilisées au collégial (Boisvert et Paradis, 2008; Boisvert, 2007; Fortier, 2003; Fréchette et Mignault, 2010; Lussier, 2004). Du côté des enseignantes et des enseignants, le constat d'une nécessité d'adapter les méthodes pédagogiques afin de mieux répondre aux besoins de cette population étudiante amène certaines questions et recherches de solution (Fréchette et Mignault, 2010; Lussier, 2004). Il y a donc un réel besoin de données pour alimenter la réflexion sur les approches et méthodes pédagogiques appropriées pour les cours de première année du DEC en sc. humaines.

Ainsi, l'ensemble de ces éléments nous amène à formuler la question générale de recherche suivante : quelles sont les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'évaluation en situation authentique dans des cours de première année du programme préuniversitaire en sciences humaines au collégial ?

DEUXIÈME CHAPITRE LE CADRE DE RÉFÉRENCE

Ce chapitre sera entièrement consacré au concept d'évaluation en situation authentique (ÉSA), le concept central de la présente démarche de recherche. L'origine et la généalogie de ce concept seront d'abord présentées, en abordant les causes et le contexte de son émergence ainsi que sa filiation avec le courant de l'évaluation des performances. S'en suit une analyse critique de la littérature qui lui est consacrée, dont les problèmes sémantiques et de définition et les différences entre les corpus francophone et anglophone. Ensuite, les caractéristiques d'une ÉSA à travers deux modèles élaborés par des auteurs et des autrices clés de ce thème sont explicitées. Finalement, une recension des problèmes reliés à la mise en œuvre rapportés dans la littérature est présentée.

1. ORIGINE ET GÉNÉALOGIE DU CONCEPT D'ÉVALUATION EN SITUATION AUTHENTIQUE

Le concept d'évaluation en situation authentique («*authentic assessment*» dans la littérature anglophone) s'est développé à partir de la fin des années 1980, principalement à travers les travaux de chercheurs américains, dont les incontournables Wiggins, Newmann et Archbald (Cumming et Maxwell, 1999; Frey, Schmitt et Allen, 2012; Palm, 2008). Le concept d'évaluation en situation authentique s'inscrit dans le courant de l'évaluation des performances («*performance assessment*»), lequel est ancré dans une critique de la validité et de la pertinence pédagogique des évaluations standardisées de type «questions à choix de réponse» (Frey *et al.*, 2012; Newmann, Brandt et Wiggins, 1998; Palm, 2008; Wiggins, 1993a).

1.1. La remise en question de l'approche psychométrique en évaluation et l'émergence du courant de l'évaluation des performances

Bien qu'on retrouve des traces d'une philosophie de l'évaluation centrée sur les performances dès l'époque prémoderne (Madaus et O'Dwyer, 1999; Popham, 1993), c'est la remise en question tout au long des années 80 et 90 de l'approche psychométrique de l'évaluation qui a favorisé la consolidation du courant de l'évaluation des performances tel que nous le connaissons aujourd'hui (Clarke, Madaus, Horn, et Ramos, 2000; Gulikers, 2006; Hoepfl, 2000; Madaus et O'Dwyer, 1999; Palm, 2008; Popham, 1993; Taylor, 1994). Cette approche était de plus en plus critiquée quant à sa capacité à évaluer les apprentissages importants et non les connaissances superficielles :

The limitations of traditional standardized tests have been well documented. They determine what is taught, they narrow content, they reduce all learning to multiple-choice format, and their results are often misused. For Worthen (1993), traditional standardized tests “do not provide insight into how students would perform on truly ‘worthy’ intellectual tasks” (p. 445). Bracey (1993) quotes a study that “concluded that commercial standardized tests and those accompanying textbooks were well constructed instruments that measure nothing particularly important” (p. 9). (Ewing, 1998,p.2)

Les découvertes importantes en sciences cognitives et l'émergence des paradigmes constructiviste et socioconstructiviste durant cette période ont accéléré le besoin pour des modèles d'évaluation alternatifs plus adaptés aux nouvelles approches pédagogiques et aux défis que posent les transformations de la société et du monde du travail pour le système d'éducation (Frey *et al.*, 2012; Newmann *et al.*, 1998; Palm, 2008; Wiggins, 1993a).

The growing interest in performance assessment and the new focus on more theoretical subjects seem to emanate from dissatisfaction with the extensive use of multiple choice tests in the US. The validity of these tests as indicators of complex performance was experienced to be too low, and to have negative effects on teaching and learning (Kane, Crooks et Cohen,

1999; Kirst, 1991). When arguing for other forms of assessment better fulfilling these requirements the term performance assessment was recognized as a suitable choice. But desires for change open up numbers of possible perspectives, so new views on the meaning of the attribute 'performance' have been added, and consensus on the meaning of performance assessment has not been reached. The dissatisfaction with the emphasis on multiple-choice testing in the US was also a fundamental factor for the development of the concept of authentic assessment. This much more recent term in education arose from the urge to meet needs that were experienced not to be met by the use of multiple-choice tests. Norm-referenced standardized multiple-choice tests of intellectual achievement were said not to measure important competence needed in life beyond school. Interpretations of test results from such tests were claimed to be invalid indicators of genuine intellectual achievement and since assessments influence teaching and learning they were also said to be directly harmful (Archbald and Newmann, 1988; Wiggins, 1989). (Palm, 2008, p. 2)

Les tenants de l'évaluation des performances proposent donc de remplacer l'approche psychométrique de l'évaluation par une démarche critériée organisée autour de la réalisation par les étudiants et étudiantes d'une tâche menant à une performance ou une production (*performance task*) (Cumming et Maxwell, 1999; Gulikers, 2006; Hoepfl, 2000; Meyer, 1992; Palm, 2008; Stecher, 2010; Wiggins, 1989*b*, 1990, 1993*a*, 1993*b*). Stecher (2010) résume ainsi ce type de tâche :

A performance task is a structured situation in which stimulus materials and a request for information or action are presented to an individual, who generates a response that can be rated for quality using explicit standards. The standards may apply to the final product or to the process of creating it. A performance assessment is a collection of performance tasks. (p. 3)

Ainsi, dans une démarche évaluative fondée sur les principes de l'évaluation des performances, les étudiantes et étudiants démontrent l'atteinte des éléments de compétence à travers la réalisation d'une tâche au cours de laquelle l'ensemble de la performance ou de la production (démarches et résultats) est pris en compte par l'évaluateur; le terme *performance* est donc utilisé «dans le sens d'*accomplissement efficace d'une tâche ou d'une opération en utilisant un ensemble intégré de*

connaissances (déclaratives, procédurales et conditionnelles).» (Louis, 2004, p.80)
 Cette tâche doit être contextualisée et complexe, exiger de la part des étudiantes et étudiants un engagement significatif et mobiliser des habiletés cognitives de haut niveau (Cumming et Maxwell, 1999; Frey *et al.*, 2012; Meyer, 1992; Palm, 2008; Stecher, 2010; Wiggins, 1998).

1.2. De l'évaluation des performances à l'évaluation en situation authentique

Parmi les tenants du courant de l'évaluation des performances, certaines autrices et certains auteurs ont formulé des critiques additionnelles vis-à-vis de l'approche psychométrique de l'évaluation, qui ont mené à l'élaboration de formes spécifiques d'évaluation des performances. C'est notamment le cas des auteurs qui ont élaboré des modèles d'évaluation en situation authentique (Archbald et Newmann, 1988; Herrington et Oliver, 2000; Schell, 2000; Wiggins, 1993*a*) qui, en plus des problèmes recensés par les tenants de l'évaluation des performances, se sont intéressés tout particulièrement aux problèmes d'arrimage entre les apprentissages scolaires et les exigences du monde non académique, notamment l'univers du travail :

The major driving force behind the need for authentic assessment is the perceived gap between what is taught, learnt and assessed in school and what is needed at work (Biemans, Nieuwenhuis, Poell, Mulder, & Wesselink, 2004; Boud, 1991). This gap arose from changes in the labour market that resulted in the rise of new kinds of jobs with different requirements for employees. In reaction, education began to change, because students needed to learn different 'things' in a different 'way' to keep up with the changes in the labour market. However, changing student learning requires new methods of assessment as assessment is possibly the most salient variable to influence student learning. These changes typify three developments that led to the need for authentic assessment. (Gulikers, 2006, p. 10)

Plus largement, de nouvelles réalités sociales entourant le savoir posent aussi des défis au monde de l'éducation, qui interpellent ces chercheurs qui ont élaboré les principes de l'ÉSA:

Today's information age is characterised by rapidly increasing information, major technological developments, and the globalisation of economies. Knowledge is now seen as a subjective and constantly changing construct that has to be created by every individual anew (Sternberg, 1999). Furthermore, knowledge and learning are seen as context-dependent. Meaningful learning requires the application of knowledge and skills in a realistic context (e.g., Brown, Collins and Duguid, 1989; Dall'Alba and Sandberg, 1996). This information society requires knowledge workers (Bereiter, 2002) and competent employees (Birenbaum, 1996) who are independent, creative and flexible problem solvers who can use, apply and develop knowledge and who can acquire the skills needed to deal with new problems within their work. In reaction to these societal changes, education began to change. (Gulikers, 2006, p. 10)

Pour pallier ces problèmes, les pionniers de l'ÉSA, tels Wiggins, Newmann et Archbald ont proposé une forme d'évaluation des performances dans laquelle le caractère réaliste des situations d'apprentissage et d'évaluation est une préoccupation de premier plan : on cherche à reproduire dans le milieu scolaire des situations d'apprentissage et d'évaluation le plus près possible des exigences que les étudiantes et étudiants sont susceptibles de rencontrer à l'extérieur de l'univers académique (Archbald et Newmann, 1988; Newmann et Archbald, 1992; Newmann *et al.*, 1998; Scheurman et Newmann, 1998; Wiggins, 1989*a*, 1989*b*, 1993*a*, 1993*b*, 1998). Cette nouvelle approche de l'évaluation, après quelques tâtonnements lexicologiques durant lesquelles les termes «*authentic achievement*», «*authentic performance*» et «*authentic task*» ont coexisté dans les écrits, s'est formalisé autour du terme d'«*authentic assessment*» dans la littérature anglophone, puis d'«*évaluation en situation authentique*» (ÉSA) dans la littérature francophone.

C'est donc dans l'effervescence entourant l'émergence du courant de l'évaluation des performances que s'est développée l'idée d'une forme d'évaluation organisée autour de tâches et de contextes les plus authentiques possible. Mais malgré cette filiation, il importe de bien distinguer les deux concepts :

Performance assessment refers to the kind of student response to be examine; authentic assessment refers to the context in which that response is performed. While not all performance assessments are authentic, it is difficult to imagine an authentic assessment that would not also be a performance assessment. [...] In a performance assessment, the student completes or demonstrates the same behavior that the assessor desires to measure. There is a minimal degree, if any, of inference involved. For example, if the behavior to be measured is writing, the student writes. The student does not complete multiple-choice about sentences and paragraphs, which instead measure the student's ability to proofread other people's writing, and require a high degree of inference about the student's ability to write. In an authentic assessment, the student not only completes or demonstrates the desired behavior, but also does it in a real-life context. (Meyer, 1992, p. 40)

Ainsi, sur le plan conceptuel et théorique, l'évaluation en situation authentique est une forme d'évaluation des performances dans laquelle le caractère authentique de la tâche proposée et de son cadre de réalisation est accentué pour devenir une préoccupation de premier plan tout au long de la démarche évaluative (Custer, 2000; Hoepfl, 2000; Meyer, 1992; Palm, 2008; Wiggins, 1998).

2. LE CONCEPT D'ÉSA DANS LA LITTÉRATURE

Une fois l'historique et la généalogie du concept d'ÉSA établi, il importe de se pencher sur sa définition précise. Qu'en dit la littérature ? Quiconque s'intéresse à la littérature à propos de l'ÉSA sera rapidement confronté à trois problèmes centraux, que nous présenterons ici : la polysémie du terme, l'absence de définition claire chez plusieurs des praticiennes et praticiens ayant présenté leur expérience dans des articles et la couverture trop superficielle du concept par la littérature francophone.

2.1. Problèmes sémantiques et de définition

Dès les premières lectures concernant l'ÉSA, on constate certains problèmes sémantiques et de définition, qui concernent la littérature francophone comme

anglophone. En effet, bien qu'un certain nombre de caractéristiques soient communes aux différentes formulations proposées par les autrices et les auteurs (dont celles synthétisées en 2.1.1), la définition du concept d'ÉSA n'est pas homogène dans la littérature (Cumming et Maxwell, 1999; Frey *et al.*, 2012; Gulikers, 2006; Meyer, 1992; Palm, 2008), tout comme c'est le cas d'ailleurs du concept de *performance assessment*:

Performance assessment and authentic assessment are recurrent terms in the literature on education and educational research. They have both been given a number of different meanings and unclear definitions and are in some publications not defined at all. Such uncertainty of meaning causes difficulties in interpretation and communication and can cause clouded or misleading research conclusions. (Palm, 2008, p. 1)

Ce flou conceptuel pose des problèmes d'analyse de la littérature théorique et d'interprétation des données portant sur des expériences d'implantation d'ÉSA. Sur le plan des écrits théoriques, toutes les autrices et tous les auteurs n'ont pas la même conception de ce qui constitue une situation authentique, et il existe différentes orientations (que nous présenterons à la section suivante) quant aux indicateurs de l'authenticité d'une situation d'évaluation à utiliser, donnant lieu à des modèles théoriques d'ÉSA présentant certaines variations. Ces divergences proviennent parfois de considérations théoriques :

The question that remains is, what is authenticity? Different researchers have different opinions about authenticity. Some see authentic assessment as a synonym to performance assessment (Hart, 1994; Torrance, 1995), while others argue that authentic assessment puts a special emphasis on the realistic value of the task and the context (Herrington & Herrington, 1998). Reeves and Okey (1996) point out that the crucial difference between performance assessment and authentic assessment is the degree of *fidelity* of the task and the conditions under which the performance would normally occur. Authentic assessment focuses on high fidelity, whereas this is not as important an issue in performance assessment. These distinctions between performance and authentic assessment indicate that every authentic assessment is a performance assessment, but not vice versa (Meyer, 1992). Messick (1994) focuses our attention to the fundamental

ambiguity that pervades all authentic assessment practices, namely, authentic to what? Honebein, Duffy and Fishman (1993) strengthen the importance of this question by saying that authenticity is a relative concept. In other words, the authenticity of something can only be defined in relation to something else. A test can either be authentic to the school or to the real world. For example, an assessment task can be authentic with respect to school problems, but inauthentic with respect to everyday life experience, because school problems do not relate to everyday life. (Gulikers, 2006, p. 22)

D'autres parts, certaines différences émergent davantage de considérations didactiques et logistiques, comme le souligne Meyer (1992):

Moreover, because authenticity has a multidimensional nature, some assessments are more authentic than others. Ironically, the most authentic assessment in many situations can probably not be contrived for purposes of testing, for then it would no longer be totally authentic. Educators and assessors must thus be explicit about which facets of authenticity are most critical. (p.40)

Du côté du corpus présentant des résultats d'expérimentation et d'implantation d'ÉSA, il n'y a pas toujours d'identification claire du modèle théorique d'ÉSA utilisé, ce qui rend le bilan de ces expériences parfois flou et difficile à évaluer sur le plan conceptuel et méthodologique. Comme le résumant bien Frey *et al.* (2012) à la suite de l'analyse de 109 publications académiques rédigées par 100 autrices et auteurs différents traitant de l'évaluation en situation authentique « *There is consensus in the field of education that classroom assessments should be authentic, but there is no consensus of what that means.* » (p.13) Cette situation impose donc une clarification des différents modèles théoriques d'ÉSA, qui sera présentée à la section 3.

2.2. Spécificités des corpus anglophone et francophone

Jusqu'à maintenant, nous avons principalement traité de la littérature anglophone abordant l'«*authentic assesement*». Qu'en est-il de la littérature francophone ? Ils existent plusieurs textes rédigés en français, principalement par des

autrices et auteurs québécois, abordant le concept d'évaluation en situation authentique et ses caractéristiques (Doirion, Merci, Pineault et Pineau, 2004; Durand et Chouinard, 2012; Guy *et al.*, 2004; Jalbert, Munn et Dodier, 2001; Legendre, 2005; Louis, 2004; Lussier et Allaire, 2004; Prigent *et al.*, 2009; Tardif, 1993; Vanasse, 2010). Toutefois, ce corpus présente des problématiques spécifiques, dont les enjeux de traduction et la nature des textes qui le composent.

On remarque d'abord en comparant les années de publication des textes en langue français et en langue anglaise que l'introduction de ce concept dans l'univers francophone s'est faite plus tardivement que dans la littérature anglophone. La première considération à propos du corpus francophone en fut donc une de traduction, étant donné que les éléments de base du concept avaient déjà été formulés par des autrices et auteurs anglophones : après quelques tentatives de traduction, c'est le terme «évaluation en situation authentique» qui s'est imposé dans la littérature francophone comme traduction du terme anglais «*authentic assessment*». Toutefois, encore à ce jour, l'uniformisation n'est pas totale (on retrouve encore une utilisation ambiguë des termes «situation authentique» et «tâche authentique» utilisés comme synonyme d'«évaluation en situation authentique»). Précisons donc que peu importe le terme utilisé dans un texte en particulier, dans le cadre de cet essai, nous utiliserons toujours le terme d'«évaluation en situation authentique», ou son abréviation ÉSA, en remplacement des termes alternatifs dans la mesure où ceux-ci réfèrent bien ce qu'est une ÉSA. Seules les citations utiliseront les termes d'origines si ceux-ci diffèrent.

Outre les questions de traduction, ce corpus se démarque aussi de la littérature anglophone par la nature des textes qui le composent. Alors que le concept d'ÉSA a émergé principalement dans des textes de nature théorique, le corpus francophone est principalement composé de textes de vulgarisation destinés aux praticiens dont l'objectif est d'offrir un survol rapide et facilement opérationnalisable sur le terrain (plutôt que d'une présentation théorique exhaustive) (Doirion *et al.*, 2004; Duval et Pagé, 2013; Guy *et al.*, 2004; Jalbert *et al.*, 2001; Lussier et Allaire, 2004; Prigent *et*

al., 2009) ou de textes présentant les résultats d'une expérience d'implantation d'ÉSA par une praticienne ou un praticien dans un contexte francophone (Vanasse, 2010). On retrouve aussi la mention du concept d'ÉSA dans des ouvrages généraux sur l'évaluation, mais il ne s'agit souvent que d'un très bref survol (Durand et Chouinard, 2012; Louis, 2004; Tardif, 1993). Aucun texte théorique approfondi traitant de façon originale de l'ÉSA n'a été recensé parmi la littérature francophone : sur le plan théorique, les textes ne font que reprendre les modèles proposés dans la littérature anglophone, et le modèle de Wiggins est nettement dominant.

Cela cause un problème important pour la recherche francophone à propos de l'ÉSA: sur le plan théorique et conceptuel, la littérature francophone offre un traitement beaucoup plus superficiel que la littérature anglophone. Notre analyse nous a permis d'identifier trois faiblesses du corpus francophone abordant le concept d'évaluation en situation authentique : 1) le degré de précision sur le plan des définitions conceptuelles est insuffisant, avec pour conséquence de présenter des assises théoriques peu approfondies; 2) les sources théoriques utilisées ne sont pas mentionnées, difficilement retraçables ou peu diversifiées, et 3) les choix théoriques quant au modèle d'ÉSA privilégié ne sont pas suffisamment précisés et justifiés. Ainsi, parce que les textes en langue française sont principalement des textes de vulgarisation destinés aux praticiennes et praticiens dont il s'agit du premier contact avec cette notion, ces textes ne sont pas adéquats pour constituer le fondement d'un cadre de référence dans une démarche scientifique (bien qu'ils demeurent très pertinents pour l'usage auquel ils sont destinés). À la suite de ce constat, nous avons choisi de fonder notre cadre de référence essentiellement sur les textes du corpus anglophones, notamment les textes fondateurs à propos de l'ÉSA rédigés par Archbald et Newmann (Archbald, 1991; Newmann et Archbald, 1992; Newmann et Wehlage, 1993) et par Wiggins (1989*a*, 1989*b*, 1990, 1993*b*, 1998), en utilisant au besoin certaines sources secondaires pour valider notre interprétation et compléter notre présentation.

3. QU'EST-CE QU'UNE ÉVALUATION EN SITUATION AUTHENTIQUE ?

Parmi les propositions de définition et de modèle d'ÉSA présents dans la littérature, nous en avons retenu deux que nous présenterons dans le détail. Le premier est celui d'Archbald et Newmann : il s'agit de la première présentation complète d'un modèle opérationnalisable d'ÉSA. Le second est celui de Wiggins, publié un an après celui d'Archbald et Newmann, qui est aussi très complet et facilement opérationnalisable pour les praticiennes et praticiens de l'éducation. C'est aussi le plus utilisé dans la littérature. Le choix de ces deux modèles repose sur trois considérations : le niveau de précision des modèles, leurs aspects opérationnalisables dans la pratique et leur caractère incontournable dans la littérature.

3.1. Le modèle de Newmann et Archbald ¹

Premiers auteurs à avoir formalisé le concept d'ÉSA sous le vocable «*authentic academic achievement*» (qu'ils utiliseront en alternance avec le terme «*authentic assessment*», dans la foulée des travaux de Wiggins) en 1988, Archbald et Newmann ont proposé un modèle d'ÉSA structuré autour de trois dimensions qui, outre quelques modifications et ajouts mineurs, est resté plutôt stable dans leurs écrits subséquents (Archbald, 1991; Newmann et Archbald, 1992; Newmann, Marks et Gamoran, 1996; Newmann et Wehlage, 1993; Newmann, 1997; Scheurman et Newmann, 1998).

Pour présenter leur modèle d'ÉSA, Archbald et Newmann misent principalement sur une présentation précise des éléments le composant, et mettent peu d'emphasis à offrir une définition synthétique du concept au lecteur. Parmi tous leurs écrits, voici l'extrait qui offre la perspective la plus globale de leur vision :

¹ Il importe de préciser que certains textes de ces auteurs ont été réalisés avec d'autres collaborateurs, comme Wehlage et Scheurman. Toutefois, comme ces textes ont été réalisés ultérieurement et qu'ils utilisent le modèle mis en place par Archbald et Newmann, nous avons choisi d'attribuer à ces derniers la paternité de ce modèle.

To summarize, this vision of authentic achievement requires students to engage in disciplined inquiry to produce knowledge that has value in their lives beyond simply proving their competence in school. Mastery of this sort is unlikely to be demonstrated in familiar testing and grading exercises. Instead, it is more often expressed in the completion of long-term projects which result in discourse, things, and performances of interest to students, their peers and the public at large. (Newmann et Archbald, 1992, p. 75)

3.1.1. *Un modèle tridimensionnel*

Pour bien comprendre le modèle d'ÉSA selon Archbald et Newmann, il faut s'intéresser à la présentation en trois dimensions qu'ils en font. Avant d'entreprendre une synthèse de celles-ci, mentionnons qu'un des attraits de ce modèle pour notre démarche est que ces auteurs ont un intérêt particulier pour les spécificités d'une ÉSA dans le domaine des sciences sociales : pour deux des trois dimensions, ils ajoutent des précisions spécifiques à ce contexte, que nous incluons dans la présentation ci-dessous.

3.1.1.1 *1^{re} dimension : La production originale de connaissance*². Pour Archbald et Newmann, une ÉSA a pour objectif premier de produire des connaissances propres à la personne la réalisant, plutôt que de simplement reproduire ce qui a été fait par d'autres en résumant sans bonification leurs travaux. Ainsi, les apprenantes et apprenants doivent réaliser leur propre production de savoir afin de construire leur compréhension des notions et habiletés mobilisées par la tâche. Cette démarche implique non seulement la compréhension des savoirs produits antérieurement par d'autres, mais également une appropriation, une reformulation et un approfondissement personnel de ces savoirs.

² Ces auteurs utilisent parfois le terme de «*production of knowledge*» et parfois celui de «*construction of knowledge*», selon les écrits.

Dans un texte subséquent (1997) visant à opérationnaliser le concept d'ÉSA, Newmann ajoute certaines précisions concernant la mise en œuvre de cette dimension dans le contexte des sciences sociales en mentionnant deux critères pour l'élaboration d'une tâche authentique dans ces domaines de la connaissance. Ainsi, une ÉSA en sciences sociales doit poser des exigences en matière

- a) d'organisation de l'information («*organization of information*») : la production de l'étudiante ou de l'étudiant doit démontrer sa capacité d'évaluer, d'interpréter et de synthétiser l'information pertinente au traitement de l'objet de la tâche et à sa réalisation ;
- b) de prise en compte des différentes perspectives («*Consideration of alternatives*») : la tâche réalisée par l'étudiante ou l'étudiant doit impliquer que celui-ci aborde le thème étudié à travers différentes visions de l'enjeu, qu'il considère une diversité de perspectives, qu'il identifie les convergences et divergences présentes.

On constate que cette première dimension de l'ÉSA implique des habiletés intellectuelles de haut niveau, qualité essentielle d'une ÉSA et passage obligé pour une réelle intégration des connaissances par les apprenantes et apprenants :

Authentic academic achievement should integrate knowledge in two ways. Not only must students be challenged to understand integrated forms of knowledge, they must also be involved in the production, not simply the reproduction, of new knowledge, because this requires knowledge integration. (Archbald et Newmann, 1988, p. 3)

3.1.1.2 2e dimension : Démarche disciplinaire structurée («*disciplined inquiry*»).

Bien que les apprenantes et apprenants soient au centre de la construction de leurs savoirs, leur démarche doit obligatoirement s'inscrire dans une approche respectant les savoirs et savoir-faire et savoir-être propres à la discipline étudiée et reconnus par les spécialistes, sous peine de mener à une production vide de sens :

Although knowledge that is constructed may be more interesting to students than knowledge that is merely reproduced, this is not to say that all constructions represent significant intellectual accomplishment. For knowledge construction to be powerful, it must be grounded on a foundation of disciplined inquiry. (Scheurman et Newmann, 1998, p. 24)

Pour respecter cette dimension, une ÉSA doit inclure trois caractéristiques, selon le modèle d'Archbald et Newmann :

1. Le recours à une base de connaissances établie et reconnue : une caractéristique centrale d'une tâche dite authentique est le recours à une démarche cognitive centrée sur les savoirs et procédures propres au domaine concerné par l'apprentissage :

For new knowledge to be significant and valid, it must be based on substantive and procedural knowledge that has been accumulated through previous workers in a field who establish facts, vocabularies, concepts, theories, algorithms, and conventions for the conduct and expression of inquiry itself. (Newmann et Archbald, 1992, p. 73)

Ainsi, il ne peut y avoir de démarche d'apprentissage et d'évaluation authentique sans recours aux savoirs, savoir-faire et savoir-être propres au domaine étudié;

2. Le développement d'une compréhension en profondeur («*In-Depth Understanding*») : La tâche à la base d'une évaluation en situation authentique doit viser le développement par les apprenantes et apprenants d'une compréhension complexe et nuancée des notions utilisées, plutôt qu'un simple survol des questions abordées.

In-depth understanding requires more than knowing a lot of details about a topic. Understanding occurs as one looks for, tests, and creates relationship among pieces of knowledge that can illuminate a given problem or issue. In short, in-depth understanding involves

construction of knowledge around a reasonably focused topic.
(Newmann, 1997, p. 364).

Ainsi, le recours à une démarche disciplinaire structurée vise à favoriser un apprentissage en profondeur plutôt que superficiel;

3. L'intégration des savoirs : Newmann et Archbald prennent soin de préciser que la maîtrise préalable des connaissances antérieures à travers une démarche centrée sur l'apprentissage en profondeur n'a pas pour objectif de pouvoir étaler son érudition, comme c'est le cas dans les évaluations standardisées, mais bien de permettre leur réelle appropriation par les apprenantes et apprenants, pour que ceux-ci puissent procéder à une intégration des savoirs antérieurs et des nouveaux savoirs qu'ils élaboreront à l'intérieur de leur production.

Pour Archbald et Newmann, une ÉSA incluant la dimension de la démarche disciplinaire structurée doit respecter ces trois caractéristiques afin d'amener les étudiantes et étudiants à démontrer leur compréhension du contenu propre à la discipline ou à la profession dans laquelle s'inscrit leur démarche (concepts, idées, théories, perspectives) ainsi que leur capacité à utiliser celui-ci à l'intérieur d'une démarche respectant les standards, les méthodes et les pratiques de cette discipline ou profession. Il s'agit là d'une condition absolue de l'authenticité d'une démarche d'évaluation, à laquelle s'ajoute la nécessité d'une valeur ajoutée à la production ou performance résultant de l'ÉSA :

the ultimate point of disciplined inquiry is to move beyond knowledge that has been produced by others; that is, to assemble and interpret information, to formulate ideas, to make critiques that cannot be easily retrieved from existing knowledge base [and] the authenticity of students' academic achievement will be enhanced if the tasks on which they are evaluated approximate these attributes of disciplined inquiry. (Archbald et Newmann, 1988, p. 2)

Mais l'élaboration de connaissances véritablement nouvelles est-elle possible à tous les niveaux du parcours scolaire ? Dans le texte publié en 1992, Newmann et Archbald amènent cette précision :

In highlighting the features of disciplined inquiry, we are not suggesting that students be expected to make seminal contributions to the academic disciplines, professions, and arts. The point is that students are capable of engaging in these forms of cognitive work when the work is adapted to their levels of development, and that authenticity of their academic achievement will be enhanced if the tasks on which they are evaluated advance production of knowledge over reproduction and if the process of production approximates these attributes of disciplined inquiry. (Newmann et Archbald, 1992, p. 74)

Ainsi, c'est bien la façon unique et personnelle de traiter l'information qui est centrale dans l'ÉSA : c'est avant tout le traitement des connaissances et habiletés impliquées dans la production résultant de l'ÉSA qui doit être propre à la personne qui l'a réalisée sans obligation de produire des connaissances véritablement inédites et nouvelles.

Tout comme pour la 1^{re} dimension, on retrouve dans la littérature subséquente des précisions quant aux exigences propres à l'application de la démarche disciplinaire structurée dans une ÉSA reliée au domaine des sciences sociales. Dans l'article de Scheurman que cosigne Newmann en 1998, les auteurs ajoutent une prescription concernant la forme que doit prendre la production finale de l'ÉSA : en raison de la place prédominante de l'écrit dans ce champ de la connaissance, la maîtrise des savoirs, savoir-faire et savoir-être disciplinaires doit être démontrée à travers une communication écrite élaborée.

3.1.1.3 3e dimension : la valeur extrascolaire de la tâche («Value beyond School»). Pour être qualifiée d'authentique, la tâche proposée doit prendre son sens dans un contexte extérieur à l'univers scolaire et sa réalisation ne doit pas viser la seule atteinte de critères académiques. *«The final, and perhaps most critical, distinction between*

authentic achievement and traditional achievements in schools is that authentic achievements have aesthetic, utilitarian, or personal value apart from documenting the competence of the learner. » (Newmann et Archbald, 1992, p. 74) p. 74

Dans son texte de 1997, Newmann précise que la valeur extrascolaire d'une tâche s'établit à l'aide de deux critères concernant le fond et la forme de l'ÉSA :

1. La question étudiée doit être en lien avec une problématique extérieure à l'école et la tâche demandée doit porter sur « *a concept, problem, or issue that is similar to one they have encountered or are likely to encounter in life beyond the classroom.* » (Newmann, 1997, p. 370);
2. La production finale doit être destinée à un public extérieur à l'école: « *The task asks students to communicate their knowledge, present a product or performance, or take some action for an audience beyond the teacher, classroom, and school.* » (Newmann, 1997, p. 371)

Bref, à travers ce modèle d'ÉSA, Newmann et ses collaborateurs posent des exigences quant au type de démarche cognitive exigée (démarche disciplinaire structurée), au type de produit attendu (une production originale de connaissance, sous forme écrite dans le cas des sciences sociales) et à l'ancrage de cette réalisation (la valeur extrascolaire de la tâche) (Archbald et Newmann, 1988; Archbald, 1991; Newmann et Archbald, 1992; Newmann *et al.*, 1996; Newmann et Wehlage, 1993; Newmann, 1997; Palm, 2008; Scheurman et Newmann, 1998), ou dans leurs propres mots :

To summarize, this vision of authentic achievement requires students to engage in disciplined inquiry to produce knowledge that has value in their lives beyond simply proving their competence in school. Mastery of this sort is unlikely to be demonstrated in familiar testing and grading exercises. Instead, it is more often expressed in the completion of long-term projects which result in discourse, things, and performances of interest to students, their peers and the public at large. (Newmann et Archbald, 1992, p. 75)

3.1.2. *Le contexte de réalisation d'une ÉSA selon Archbald et Newmann*

En plus du respect des trois dimensions présentées ci-dessus, le modèle d'Archbald et Newmann prescrit certains critères quant au contexte de réalisation d'une ÉSA, dont quatre sont explicités dans le texte conjoint de Newmann et Archbald publié en 1992. Ainsi, pour qu'une activité d'apprentissage et d'évaluation soit réellement authentique, elle doit se dérouler dans un contexte qui permet :

1. La collaboration entre pairs et l'interaction des apprenantes et apprenants avec l'enseignante ou l'enseignant. Ici, le modèle de Newmann et Archbald s'inscrit clairement dans une approche de co-construction des savoirs, dans laquelle les étudiantes et étudiants sont à la fois aidants et aidés;
2. L'accès à des ressources et outils appropriés: La réalisation de l'activité doit inclure un accès aux ressources et outils appropriés à la tâche demandée normalement disponibles dans un cadre extrascolaire, et cela doit primer sur des exigences de mémorisation d'informations. Ainsi, les étudiantes et étudiants doivent être soumis à un contexte réaliste quant aux ressources qui sont normalement accessibles à une personne devant réaliser une tâche similaire dans le monde extrascolaire. Pour Newmann et Archbald :

Becoming educated in medieval times, before printing, required memorizing the contents of key manuscripts, because future access was unlikely. But today, the specialization and explosion of knowledge create a new challenge. Possession of prior knowledge is critical to disciplined inquiry, but it is now impossible to teach students all the relevant information needed to complete authentic tasks. The new challenge is to learn how to use tools and resources to find the appropriate missing knowledge. (Newmann et Archbald, 1992, p. 78)

3. Le temps alloué et son organisation sont adaptés aux caractéristiques de la tâche ainsi qu'aux besoins des étudiantes et étudiants. Le rythme de travail est constant, mais d'intensité modérée, propice à un apprentissage en profondeur. (Archbald et Newmann, 1988; Archbald, 1991; Newmann et Archbald, 1992; Newmann et Wehlage, 1993; Scheurman et Newmann, 1998);

4. Une certaine autonomie des apprenantes et apprenants quant au déroulement et à l'issue de l'activité :

Rather than always toiling under predetermined routines arbitrarily dictated by authorities, the authors of authentic work need to influence the conception, execution, and evaluation of the work itself. [...] There are, of course, important limits on the extent to which students should control the learning of academic subjects. Certain facts, definition, concepts, algorithms, and processes of verification must be assimilated according to standards in the fields of knowledge to be taught. But for kind of learning to be translated into authentic products, students must enjoy some autonomy and discretion in its application. (Newmann et Archbald, 1992, p. 79)

Ces prescriptions contextuelles reposent sur la volonté de s'approcher le plus possible du cadre extrascolaire, lequel offre la plupart du temps une flexibilité temporelle et collaborative, et d'ainsi augmenter l'authenticité de l'activité d'évaluation.

In short, authenticity will be enhanced if assessment tasks have aesthetic or utilitarian value beyond instructional evaluation, and this is most likely when the tasks involve student production, rather than reproduction, of knowledge, flexible use of time, and access to help from others in completing the task. (Archbald et Newmann, 1988, p. 4)

3.1.3. Un 5^e critère contextuel : le contexte de réalisation permet l'évolution de l'apprentissage

À ces quatre critères contextuels formellement identifiés par Archbald et Newmann, nous nous proposons d'en ajouter un qui, sans avoir été formellement identifié comme une caractéristique contextuelle distincte par ces auteurs, traverse l'ensemble de leurs textes et est suffisamment central à notre avis pour être explicitement nommé comme critère contextuel intrinsèque de l'ÉSA, soit la possibilité d'évolution de la maîtrise des connaissances et compétences durant la réalisation de la tâche. Tout au long des écrits d'Archbald et Newmann, ainsi que des textes cosignés avec d'autres collaborateurs (notamment le texte écrit par Newmann et Wehlage

(1993)), l'idée que le contexte de réalisation d'une ÉSA doit permettre l'évolution de la maîtrise des connaissances et compétences tout au long de la tâche est omniprésente : c'est d'ailleurs un des arguments invoqués pour remplacer l'approche psychométrique, où l'évaluation advient à la fin d'un cycle d'apprentissage afin de sanctionner l'atteinte ou non des cibles fixées, par l'ÉSA qui permet aux étudiantes et étudiants de perfectionner leur compréhension et leur maîtrise des éléments de compétence visés par la tâche tout au long de sa réalisation. Une des forces de l'ÉSA identifiée par les auteurs est que celle-ci est la démarche à travers laquelle l'apprentissage se réalise et la production finale évaluée est produite. (Archbald et Newmann, 1988; Archbald, 1991; Newmann et Archbald, 1992; Newmann et Wehlage, 1993; Scheurman et Newmann, 1998). Ainsi, le produit ou la performance soumis à l'évaluation sommative est le résultat d'une réelle démarche d'apprentissage, et donc un témoin/une démonstration du niveau de maîtrise des savoirs, savoir-faire et savoir-être développés. Or, pour que cela soit possible, le contexte de réalisation doit permettre cette évolution, d'où l'importance de l'inclure dans les critères contextuels de réalisation d'une ÉSA.

Pour compléter la présentation du modèle d'Archbald et Newmann, mentionnons le travail de formulation d'un outil en cinq échelles réalisé par Newmann et Wehlage (1993) – basé sur les principes et dimensions présentés jusqu'à maintenant – qui permet de mesurer et comparer le degré d'authenticité d'une ÉSA. Dans cet article, les auteurs réitèrent que l'évaluation du caractère authentique d'une évaluation ne se fait pas sur un mode binaire, mais plutôt à partir d'un continuum, avis qui est d'ailleurs partagé par plusieurs spécialistes de l'ÉSA (Cronin, 1993; Gulikers, 2006; Wiggins, 1989a, 1998).

Pour Newmann et Wehlage (1993), le degré d'authenticité d'une tâche se détermine en la positionnant sur les 5 échelles suivantes :

1. *Le recours à des habiletés intellectuelles de haut niveau* : Une ÉSA ayant un faible degré d'authenticité pour ce critère correspond à une tâche demandant aux

étudiantes et étudiants de reproduire ou réciter une notion apprise par cœur, alors qu'un haut niveau d'atteinte de cet objectif s'incarne dans une tâche exigeant une manipulation des notions et connaissances à travers une démarche propre et personnelle, de laquelle résulte un produit original;

2. *La profondeur de l'apprentissage* : Cette échelle est destinée à mesurer la qualité des apprentissages quant à la profondeur ou à la superficialité de la compréhension et à l'appropriation des connaissances acquises par les étudiantes et étudiants durant une tâche : un faible degré d'authenticité pour ce critère réfère à un apprentissage superficiel des connaissances impliquées dans la tâche alors qu'un niveau d'atteinte élevé indique que la réalisation a permis une maîtrise en profondeur et une appropriation réelle de celles-ci par les apprenantes et apprenants;
3. *Ancrage dans le monde extrascolaire* : Cette échelle permet d'évaluer le niveau de pertinence de la tâche proposée vis-à-vis du monde extrascolaire. Ainsi, une activité d'apprentissage et d'évaluation centrée sur des habiletés qui n'ont d'importance ou d'utilité que dans le milieu scolaire se positionnera comme ayant un faible degré d'authenticité pour ce critère ; à l'inverse, une tâche qui a une pertinence dans le contexte social ou professionnel dans lequel les étudiantes et étudiants évoluent sera considérée comme très authentique pour ce critère;
4. *Raisonnement en profondeur* : Pour être authentique, une activité d'apprentissage et d'évaluation doit impliquer une analyse et une discussion de fond. Ainsi, un positionnement faible sur cette échelle indique que la tâche demande peu de réflexion interactive, n'exigeant que des justifications routinières, alors qu'une activité présentant un haut niveau d'authenticité pour ce critère exigera une démarche singulière et non routinière de réflexion et d'interaction entre les notions impliquées dans la réalisation de la tâche qui contribue à un dialogue collectif constructif favorisant la meilleure compréhension des thèmes étudiés;
5. *Soutien mutuel à la réussite* : Cette échelle s'intéresse à la présence d'un soutien à la réussite inclusif et respectueux tout au long de la démarche d'apprentissage. Un contexte d'apprentissage et d'évaluation offrant un faible soutien à la réussite se caractérise par un environnement décourageant les efforts et la participation de tous,

une atmosphère négative et peu propice à l'épanouissement intellectuel. À l'inverse, un contexte où il existe un haut niveau de soutien mutuel à la réussite se caractérise par un climat propice à la réussite de tous où l'entraide permet de maximiser les forces de toutes et tous dans le respect des capacités et des contributions de chaque personne et où l'initiative est valorisée et soutenue par l'ensemble des personnes impliquées dans le processus.

3.1.4. *L'évaluation d'une ÉSA selon le modèle d'Archbald et Newmann*

Finale­ment, dans un article publié en (1991), Archbald précise certains principes évaluatifs indissociables d'une ÉSA. Ainsi, l'évaluation de la performance ou production résultant d'une démarche conforme au modèle d'Archbald et Newmann doit

- a) *reposer sur des indicateurs multiples* : les différentes dimensions des connaissances et habiletés exigées au long de l'activité d'évaluation doivent être prises en considération au moment de porter un jugement sur la qualité du produit ou de la performance résultant de la démarche;
- b) *utiliser un barème d'évaluation critérié* : les outils d'évaluation sont construits à partir de normes de performance préétablies permettant de qualifier le niveau de compétences démontré pour chaque indicateur en regard des exigences propre à la tâche demandée, indépendamment de la performance des autres étudiantes et étudiants du groupe;
- c) *faire appel au jugement professionnel de personnes qualifiées* : l'évaluation d'une production ou performance résultats d'une démarche complexe et authentique nécessite le discernement d'un regard expert habileté à utiliser les barèmes d'évaluation associés aux indicateurs impliqués dans une tâche complexe et d'en faire une interprétation globale et non simplement comptable (comme c'est le cas

des tests standardisés). Archbald suggère également l'implication de plusieurs évaluatrices et évaluateurs

Tasks that are complex and judgments that have significant consequences often require more than one assessor. It is crucial, of course, that indicators and criteria of quality are agreed upon. Achieving high interrater reliability is necessary to establish the legitimacy in the eyes of students and other audiences of the practice of authentic assessment. p. 283-284 (Archbald, 1991, p. 283-284)

On constate que sur le plan des principes évaluatifs, l'ÉSA s'inscrit parfaitement dans l'approche de l'évaluation des performances (Meyer, 1992; Newmann *et al.*, 1996; Palm, 2008; Stecher, 2010).

3.2. Le modèle d'ÉSA selon Wiggins :

Grant Wiggins est un autre auteur central du développement du concept d'ÉSA et a grandement contribué à la définition et à la popularisation du concept d'ÉSA (Ashford-Rowe, Herrington et Brown, 2014; Custer, 2000; Frey *et al.*, 2012; Meyer, 1992; Palm, 2008), et il est celui qui domine la littérature récente, tout particulièrement la littérature francophone (Doirion *et al.*, 2004; Duval et Pagé, 2013; Guy, s.d.; Jalbert *et al.*, 2001; Louis, 2004; Lussier et Allaire, 2004; Prigent *et al.*, 2009).

Les travaux de Wiggins s'inscrivent dans une volonté de revoir le paradigme psychométrique pour favoriser une approche de l'évaluation davantage intégrée à l'enseignement et au monde extrascolaire. Pour cet auteur, l'approche traditionnelle de l'enseignement et de l'évaluation est complètement déconnectée de la réalité («*real world*»), et ce fonctionnement en vase clos ne permet pas aux étudiantes et étudiants de développer des compétences qui leur seront utiles dans leur vie professionnelle, sociale et civique (Wiggins, 1989a, 1989b, 1990, 1998). Il est donc impératif de délaisser cette approche de l'enseignement et de l'évaluation qui mise sur le

développement d'habileté utile essentiellement dans l'univers académique pour lui substituer une approche plus authentique :

Authentic assessment is simply a fancy way of describing what performers in the world beyond school face each day: they are tested on their ability to do a known job well and in a variety of circumstances in which the performer has a right to receive or seek utter clarity about purposes, audience, standards, and criteria; they are also quizzed or challenged to cope with isolated facts, decisions, or needs; but there is invariably nothing unknown about the purpose of the challenges they face, about how to prepare them, and about how to link them to larger objectives. They get better at their job because they know what is expected of them, and they can check feedback from bosses or clients that can be sought at any time. Why should school testing be any different ? (Wiggins, 1998, p. 4)

Mais comment intégrer concrètement cette approche dans des activités d'apprentissage et d'évaluation ? À travers ses écrits s'échelonnant sur une décennie, Wiggins forge un modèle d'ÉSA qui lui est propre.

3.2.1. *La définition de l'ÉSA et ses caractéristiques fondamentales selon Wiggins*

En 1989, dans son premier article consacré principalement à l'ÉSA, il propose la définition suivante :

In sum, authentic tests have four basic characteristics in common. First, they are designed to be truly representative of performance in the field; only then are the problems of scoring reliability and logistics of testing considered. Second, far greater attention is paid to the teaching and learning of the criteria to be used in the assessment. Third, self-assessment plays a much greater role than in conventional testing. And fourth, the students are often expected to present their work and defend themselves publicly and orally to ensure that their apparent mastery is genuine. (Wiggins, 1989b, p. 45)

En 1998, dans son ouvrage le plus complet (1998) consacré au principe d'authenticité en enseignement et en évaluation, Wiggins présente 6 principes

définissant une ÉSA. Pour Wiggins (1998), une activité d'évaluation ou d'apprentissage est authentique si

- a) *elle est réaliste*, c'est-à-dire si elle vise un niveau de mobilisation et de maîtrise des connaissances et habiletés évaluées comparables aux exigences rencontrées dans des situations réelles du monde extrascolaire;
- b) *elle requiert le recours au jugement et à l'innovation*, c'est-à-dire qu'elle nécessite une utilisation judicieuse de connaissances et habiletés appropriées pour résoudre un problème non structuré nécessitant la création d'une solution non convenue. Une tâche authentique exige donc davantage que le simple recours à une procédure routinière ou des connaissances réponses réflexes (*«plug-in knowledge»*);
- c) *elle demande une réalisation originale*, propre à chaque participant, c'est-à-dire que les étudiantes et étudiants doivent aller au-delà de la simple reproduction de ce qu'on leur a enseigné en «se plongeant» dans l'objet d'étude à travers la réalisation d'une tâche reproduisant l'approche de la discipline étudiée, mais menant à un résultat inédit;
- d) *elle reproduit le plus fidèlement possible le contexte dans lequel les adultes sont réellement «évalués» dans le monde du travail, la vie civique et la vie personnelle*, c'est-à-dire que la tâche demandée doit permettre d'expérimenter les contraintes et caractéristiques de situations réelles existant véritablement à l'extérieur de l'école, lesquelles ne sont pas toujours aussi structurées et ordonnées qu'un problème théorique;
- e) *Elle évalue la capacité d'utiliser efficacement et adéquatement un corpus de connaissances et d'habiletés afin d'effectuer une tâche complexe*, c'est-à-dire que l'activité d'apprentissage et d'évaluation permet de vérifier l'autonomie dans le choix des outils et connaissances à mobiliser pour faire face adéquatement à un problème complexe ainsi que sa capacité à les utiliser adéquatement;
- f) *Elle offre la latitude nécessaire pour consulter des ressources, obtenir du soutien et de la rétroaction, et effectuer les étapes permettant de développer, perfectionner et améliorer les performances et résultats*, c'est-à-dire que l'organisation et le

déroulement de la tâche permettent aux étudiantes et étudiants de retravailler les différents aspects de la réalisation, de faire des essais et erreurs, de consulter des ressources et d'obtenir des rétroactions sur leur progression, de façon à développer leur maîtrise des connaissances et habiletés impliquées dans la tâche et de peaufiner la performance ou le produit final.

Ces éléments sont les grands principes sur lesquels repose l'ÉSA selon le modèle de Wiggins.

3.2.2. *Un modèle quadridimensionnel à l'usage des praticiennes et praticiens*

À travers les chapitres subséquents de son ouvrage de 1998 et dans ses écrits antérieurs (1989a, 1989b, 1993a, 1993b, 1998), Wiggins aborde plus en détail les aspects concrets de la mise en application de ces principes et ajoute des précisions, de façon à en faciliter l'opérationnalisation et l'usage par les praticiennes et praticiens de l'éducation. Dans le but de présenter une synthèse complète et rigoureuse de ce modèle, nous avons regroupé tous les éléments et précisions présentées par Wiggins au fil de ses publications dans une synthèse organisée autour de quatre dimensions (un découpage utilisé par Wiggins dans deux de ses textes fondateurs (Wiggins, 1989a, 1989b)), qui résume le modèle de l'ÉSA selon Wiggins :

1.2.2.1 *1^{re} dimension : Structure et logistique («Structure and logistics»)*. Cette dimension est composée des éléments suivants :

1. Comporte plusieurs étapes ou plusieurs moments de réalisation (contrairement au modèle de l'examen final) dont les composantes et le déroulement doivent être connus à l'avance;
2. Les composantes et le déroulement sont connus à l'avance, idéalement dès le début de la tâche;

3. Les exigences et caractéristiques de la tâche dictent l'organisation pédagogique (et non l'inverse) et aucune contrainte de temps arbitraire ou irréaliste n'est imposée;
4. La performance ou production finale est destinée à un public (idéalement extrascolaire);
5. Implique une dimension de collaboration avec autrui (enseignantes et enseignants, collaborateurs, autres étudiantes ou étudiants);
6. L'évaluation sommative implique des critères multiples connus à l'avance, et idéalement plusieurs évaluateurs;
7. Est itératif et permet aux étudiantes et étudiants de pratiquer les habiletés à acquérir et perfectionner les savoirs à développer, objectif ultime de la démarche;
8. Offre une rétroaction constante et l'évaluation formative tient une place prépondérante durant la démarche.

3.2.2.1. 2^e dimension : *Composantes intellectuelles de la tâche («Intellectual design features»)*. Cette dimension est composée des éléments suivants :

1. Est organisé autour d'une tâche contextualisée, représentative des défis réellement rencontrés dans la vie extrascolaire, sur le plan professionnel, personnel ou civique;
2. Est un défi, complexe, significatif, multidimensionnel et intégrateur (et non une série des tâches atomisées);
3. Porte sur les éléments essentiels des habiletés et savoirs à développer et non sur des éléments pointus et périphériques choisis arbitrairement ou sur la simple reproduction d'information;
4. Axé sur le développement en profondeur des habiletés et connaissances;
5. L'évaluation n'est pas une fin en soi, mais vise l'amélioration à long terme de la maîtrise des connaissances et habiletés impliquées dans la tâche : l'évaluation du résultat n'est pas un but en soi, mais un moyen pour permettre aux étudiantes et étudiants d'améliorer de façon autonome leurs performances futures grâce à la rétroaction et l'autorégulation des apprentissages;

6. Contient un certain degré d'ambiguïté visant à refléter les défis réels qui sont rarement aussi bien définis que des problèmes théoriques;
7. Mène à une réalisation originale, développée de façon autonome et unique par chaque étudiant (ou groupe d'étudiants);
8. Comprend une justification par l'étudiante ou l'étudiant des démarches et des choix effectués pour parvenir à la réalisation finale, exposant ainsi les stratégies cognitives et métacognitives de l'étudiante ou l'étudiant;
9. Est conçu comme un processus intégré à l'apprentissage (et non une évaluation séparée de la démarche d'apprentissage) et permet à l'étudiante ou l'étudiant de développer et perfectionner sa maîtrise des habiletés et des connaissances évaluées tout au long de la réalisation de la tâche;
10. Doit être perçue comme intéressant, approprié et significatif sur le plan éducatif par les étudiantes et étudiants comme par les enseignantes et enseignants et doit tenir compte de l'intérêt, des capacités et du niveau de persévérance des étudiantes et des étudiants;
11. Vise l'autonomie des étudiantes et étudiants, autant sur le plan de l'identification des connaissances et habiletés à mobiliser, de la manière de le faire que de l'autoévaluation et l'autorégulation de ses pratiques.

3.2.2.2.3^e dimension : *Critères d'évaluation et de notation*. Cette dimension est composée des éléments suivants :

1. Est une évaluation critériée³, reposant sur des standards de performance reconnus et préétablis, et non à partir d'une comparaison avec le groupe;
2. Évalue les performances et non les efforts ou les intentions;
3. Les critères d'évaluation ciblent les éléments centraux de la discipline, des habiletés et des connaissances propres à la tâche demandée (plutôt que ce qui est facilement évaluable, mais pas toujours central);
4. Permet de distinguer si une compréhension et une maîtrise véritables (plutôt qu'un apprentissage superficiel relevant de la simple familiarité avec le sujet) ont été atteintes;
5. L'évaluation sommative porte uniquement sur la performance ou production finale, et non sur les étapes antérieures durant lesquelles l'apprentissage n'était pas complété;
6. Les critères d'évaluation sont connus à l'avance (le plus tôt possible dans le processus), faciles à comprendre et sont considérés comme légitimes par les étudiantes et étudiants;
7. Utilise un système de notation multidimensionnel plutôt que l'attribution d'une seule note globale;
8. Comporte une dimension autoévaluative;
9. S'inscrit harmonieusement dans la philosophie éducative globale de l'institution.

³ Dans son Dictionnaire actuel de l'éducation (2005), Legendre définit une évaluation critériée comme «un mode d'évaluation où la performance du sujet dans l'Accomplissement d'une tâche spécifique est jugée par rapport à un seuil ou à un critère de réussite, déterminé dans la formulation du ou des objectifs explicitement visés, indépendamment de la performance de tout autre sujet» p. 634

3.2.2.3.4^e dimension : *Équité et transparence*. Cette dimension est composée des éléments suivants :

1. Minimise les comparaisons entre les personnes évaluées;
2. Prend en compte la diversité des aptitudes, styles d'apprentissage et intérêt parmi les étudiantes et étudiants;
3. Ancre la validité et la fiabilité de l'évaluation d'abord dans la pertinence du choix des éléments évalués (et non leur uniformité) et dans leur caractère significatif vis-à-vis des normes de la vie adulte et des exigences réellement en vigueur dans un champ disciplinaire;
4. Doit être conçu de façon à ce que le cheminement fournisse à tous la possibilité réelle d'améliorer la maîtrise des habiletés et connaissances, sans jamais succomber à la tentation d'un nivellement vers le bas pour dissimuler un niveau de compétence faible ou un manque d'expérience ou de pratique;
5. Favorise l'échange entre la personne qui évalue et celle qui est évaluée, et offre la possibilité à cette dernière de fournir des précisions supplémentaires au besoin;
6. Vise le maintien d'un équilibre équitable pour toutes et tous entre les performances atteintes à travers le travail, les habiletés naturelles et les apprentissages antérieurs reliés à un contexte extrascolaire;
7. Favorise l'expression des forces des étudiantes et étudiants dans toutes leurs dimensions, plutôt que le recours à des normes strictes et prédéterminées, afin de favoriser l'épanouissement de toutes les facettes de leurs habiletés.

3.3. Comparaison des modèles d'Archbald et Newmann et de Wiggins : deux constats pertinents pour notre démarche de recherche

Bien que structurés différemment, les modèles de l'ÉSA selon Archbald et Newmann et selon Wiggins ont une large majorité de caractéristiques et critères en commun, malgré quelques nuances. Quel modèle est le plus adapté à notre démarche ? La définition de ce qu'est l'authenticité d'une démarche d'évaluation, sans amener un rejet de Wiggins, rend incontournable le modèle d'Archbald et de Newmann.

Dans les deux modèles, l'authenticité d'une évaluation se définit d'abord par sa proximité avec les exigences et attentes extérieures au contexte scolaire immédiat dans lequel se réalise l'activité. Ainsi, pour ces trois auteurs, les tâches qui composent l'activité d'évaluation soumise aux étudiantes et aux étudiants doivent être centrées sur les attentes et exigences qui existent réellement dans la «vraie vie». Par contre, Wiggins est entièrement tourné sur le monde extrascolaire, alors que le modèle d'Archbald et Newmann offre une conception plus fluide de la frontière entre le monde académique et non académique, en reconnaissant les exigences du monde académique universitaire comme des normes d'authenticité valides pour la définition des standards à atteindre par les étudiantes et étudiants des niveaux antérieurs (Newmann, 1997; Scheurman et Newmann, 1998).

[Dans le modèle d'Archbald et Newmann,] the third criterion of authenticity, that the accomplishment should have value beyond school, is also related to the desired product but could also be seen as requiring the figurative context dealing with issues that have meaning beyond school. Wiggins' perspective of authentic assessment is also 'life beyond school' and in addition to 'processes and products' he also emphasizes 'conditions'. He does not specify the 'processes and products' in the same way that Archbald and Newmann do but claims that in authentic assessment "The tasks are either replicas of or analogous to the kinds of problems faced by adult citizens and consumers or professionals in the field" (1993, p. 206), and that "replicating or simulating the diverse and rich contexts of performance" (1993, p. 207) is the most important one of his nine criteria of authenticity. (Palm, 2008, p. 7)

Ainsi, chez Archbald et Newmann, l'idée qu'une tâche authentique doit avoir une «valeur au-delà de l'école» signifie davantage que l'activité doit avoir un sens au-delà du simple exercice académique et du contexte immédiat de réalisation en classe, mais n'insiste pas autant que Wiggins sur le caractère extrascolaire du produit ou de la performance résultant de l'ÉSA. Ainsi, chez Archbald et Newmann, un problème dont la nature concerne des enjeux extérieurs à l'école, mais dont la résolution se fait avec les outils et démarches propres au monde académique peut conserver son caractère authentique, et la porte reste ouverte à ce que l'acquisition de connaissances et habiletés

nécessaires à bien performer dans un univers académique de niveau supérieur, comme l'université, soit considérée comme un indicateur d'authenticité d'une tâche. Chez Wiggins, on cherche davantage un ancrage extrascolaire à la démarche comme au produit de l'ÉSA (Wiggins, 1998) : le recours à des indicateurs centrés sur un palier académique supérieur ou le fait que l'enjeu placé au cœur de l'ÉSA ait une existence réelle à l'extérieur du contexte immédiat d'apprentissage, bien qu'il soit un pas dans la bonne direction, n'est pas suffisant pour atteindre un haut degré d'authenticité. Sur ce point, notre étude d'ÉSA dans le contexte de cours du DEC en sciences humaines, un programme dont le principal objectif est la préparation des étudiantes et des étudiants aux études universitaires, est davantage compatible avec le modèle d'Archbald et Newmann.

Cette compatibilité plus grande avec le modèle d'Archbald et Newmann se constate aussi dans les spécifications quant au type de produit attendu pour une ÉSA dans l'univers des sciences sociales : Archbald et Newmann ont précisé des conditions spécifiques au fond et à la forme attendus du produit d'une ÉSA dans les disciplines de l'univers des sciences sociales, démontrant ainsi la compatibilité de leur modèle avec un contexte semblable à celui dans lequel se réalise notre étude (Newmann, 1997; Scheurman et Newmann, 1998). Cette compatibilité plus naturelle entre l'univers des sciences humaines et le modèle d'Archbald et Newmann est d'ailleurs invoquée par Riley et Slater Stern (1998), dans leur texte présentant un bilan de la création et de l'implantation d'une ÉSA destinée à des étudiantes et étudiants en enseignement, pour expliquer leur choix d'utiliser ce modèle comme cadre théorique.

Mentionnons toutefois que certains textes présentent des analyses d'ÉSA dans des domaines tels l'histoire (McMann et Jepsen McMann, 1992) qui ont été élaborés à partir d'une interprétation flexible du modèle de Wiggins montre le potentiel de ce modèle dans l'enseignement des sciences sociales, tout en illustrant sa plus grande distance avec celui-ci en raison de la nécessité d'une interprétation plus large du modèle.

4. LES DIFFICULTÉS RENCONTRÉES DANS LA MISE EN ŒUVRE DE L'ÉSA : QU'EN DIT LA LITTÉRATURE ?

Outre la question de la définition, des caractéristiques et des critères d'une ÉSA, la littérature consultée discute abondamment des problèmes reliés à la mise en œuvre d'évaluation en situation authentique, enjeux au cœur de notre démarche dans le cadre de cet essai. Nous avons donc recensé les problèmes invoqués dans les textes de nature théorique, ainsi que dans les textes présentant des expériences d'implantation d'ÉSA par des praticiennes et praticiens, afin d'en faire une synthèse. Ces informations sont également résumées sous forme de tableau à l'annexe E (Tableau 1 – Synthèse des problèmes reliés à la conception et la mise en œuvre d'une ÉSA recensés dans la littérature).

4.1. La qualité et la rigueur du processus évaluatif :

La capacité des enseignantes et enseignants et des institutions académiques à assurer l'équité, la fiabilité et la validité des évaluations réalisées selon les principes de l'ÉSA est une préoccupation des plus présentes dans la littérature (Ewing, 1998; Gresham, 1991; Hoepfl, 2000; Kamphaus, 1991; Maeroff, 1991; Messick, 1994, 1996; Montgomery, 2002; Riley et Slater Stern, 1998; Tanner, 2001; Taylor, 1994; Terwilliger, 1997).

The consistency of the judgments that are made in assessment (reliability), and the degree to which one can demonstrate the efficacy of one's evaluations of students' performance (validity) are relevant whatever their assessment context. The instructor who evaluates the science portfolios should wonder whether another similarly prepared instructor would agree with the judgments. Where this discussion has occurred in the professional literature (see Koelsch, Estrin, and Fan, 1995), experimenters generally conclude that developing a reasonable level of interrater reliability is difficult and expensive because of the training required (Mehrens, 1992; Reckase, 1995). The alternative, however, is to allow the instructor's judgment to be a major source of assessment error. The hallmark of authentic assessment practices is their harmony with real-world circumstances. This appearance factor is very much akin to what in

traditional testing is called face validity, the appearance of validity. It is ironic that this least substantive dimension of validity would become a central tenet in authentic assessment. (Tanner, 2001, p.27-28)

Les études d'expérience d'implantation d'ÉSA recensées par Hoepfl (2000) et par Wiggins (2011) en arrivent aux mêmes conclusions : il est possible d'atteindre un haut degré de fiabilité dans l'évaluation des ÉSA, mais cela demande une formation adéquate des évaluateurs et évaluatrices et des périodes de temps consacrées à la réalisation d'exercices de concordance des jugements, ce qui nous ramène à la question des ressources (Montgomery, 2002; Wiggins, 2011).

Mais qu'en est-il lorsqu'il n'est pas possible – comme c'est le cas pour les enseignantes et enseignants du collégial qui prennent l'initiative individuellement d'implanter une ÉSA dans le cadre de leurs cours - d'augmenter les ressources et d'établir des pratiques collaboratives entre les évaluatrices et évaluateurs ? Certaines autrices et certains auteurs ont conclu qu'il fallait alors accepter le sacrifice d'un certain niveau de fiabilité au profit de la validité des évaluations elles-mêmes et des impacts largement positifs de ce mode d'évaluation sur l'apprentissage réel des étudiantes et étudiants (Montgomery, 2002; Riley et Slater Stern, 1998; Wiggins, 1990). Wiggins abonde dans ce sens, en argumentant que le maintien du statu quo en faveur du paradigme psychométrique n'est pas non plus neutre sur le plan de la validité et de la fiabilité, car en raison de la nécessité de simplifier et de décontextualiser les questions afin de les rendre décomposables en réponses objectives, les tests standardisés sacrifient la validité des évaluations au profit de la fiabilité des résultats: *«test are intrinsically prone to sacrifice validity to achieve reliability and to sacrifice the student' interests for the test maker's»* (Wiggins, 1993a, p. 4). Ainsi, pour Wiggins, le choix à privilégier est clair : étant donné que le recours aux tests standardisés est imparfait, car il sacrifie la validité des évaluations au profit de leur fiabilité et que les avantages de l'ÉSA sur le plan de l'apprentissage et de la validité des évaluations dépassent les inconvénients liés à la perte de fiabilité occasionnée par l'abandon de l'approche psychométrique des évaluations, cette dernière option doit prévaloir.

Autre aspect soulevé par certaines autrices et certains auteurs : la question du caractère généralisable ou non des résultats de l'ÉSA en tant qu'évaluation certificative. Étant donné le caractère contextualisé de l'ÉSA, l'évaluation en résultant est-elle réellement représentative des connaissances et habiletés développées par les étudiantes et les étudiants ? Peut-il y avoir des biais induits par le contexte de réalisation ? Pour Yen (1993, dans Hoepfl, 2000), la question est fondamentale. Messick (1996) accorde également une grande importance à cet enjeu :

This issue of generalizability of score inferences across tasks and contexts goes to the very heart of score meaning. Indeed, setting the boundaries of score meaning is precisely what generalizability evidence is meant to address. However, because of the extensive time required for the typical performance task, there is a conflict in performance assessment between time-intensive depth of examination and the breadth of domain coverage needed for generalizability of construct interpretation. (p. 11)

Ainsi, parce que les ÉSA ne permettent pas d'évaluer un aussi large spectre de connaissances et d'habiletés que les tests standardisés, le choix et la conception des ÉSA commandent une sélection rigoureuse afin d'assurer le caractère réellement représentatif de l'apprentissage global réalisé par l'étudiante ou l'étudiant du jugement porté par l'évaluateur ou l'évaluatrice (Brennan, 1996; Haertel et Linn, 1996; Haertel, 1999; Messick, 1994, 1996).

Les enjeux quant à la sélection des tâches incluses dans une ÉSA sont également une question d'équité pour l'ensemble des étudiantes et étudiants. Or, cet aspect est beaucoup plus complexe à évaluer qu'il ne le semble.

Un effet documenté par plusieurs auteures et auteurs d'impacts dus à la contextualisation de certaines ÉSA est la réussite différenciée selon l'origine ethnique et la langue maternelle des étudiantes et étudiants. « Koelsch, Estrin, and Fart (1995) noted: "The level of language that students are expected to process and produce in the course of completing performance assessments is nearly always more complex than the

language of traditional standardized tests. Students who are still learning English and those who have not grown up in households where language forms and uses parallel those of the classroom are likely to be at an even greater disadvantage with performance assessments than they were with multiple-choice and short answer tests.” (p. 15) Authentic mathematics assessments often require students to explain their solutions. The elaboration provides important insight regarding the students' grasp of the content, but it also taxes their language abilities. (Tanner, 2001, p. 28)

En effet, lors de certaines expériences d’implantation d’ÉSA, on a constaté que ce type d’évaluation défavorisait systématiquement les étudiantes et étudiants dont le niveau de maîtrise du langage est plus faible, ou ceux ayant un bagage culturel différent de la majorité, car les explications reliées aux démarches et les références intégrées à la contextualisation, encouragées dans l’ÉSA, entraînent une répercussion de ces caractéristiques dans l’ensemble des matières (Bond, Moss et Carr, 1996; Brennan, 1996; Haertel, 1999; Hoepfl, 2000; Messick, 1996; Montgomery, 2002; Tanner, 2001).

4.2. Les exigences en temps et en ressources

Pour plusieurs spécialistes, la quantité de ressources financières et professionnelles que requiert l’implantation d’ÉSA est le principal obstacle à la généralisation de cette approche évaluative (Ewing, 1998; Hardy, 1996; Khattri, Reeve, et Kane, 1998; Maeroff, 1991; Monk, 1996; Montgomery, 2002; Tanner, 2001; Taylor, 1994). Wiggins rétorque qu’il s’agit d’un argument non fondé, car «*the ‘test’ of performance is the course, not something you do after it* » (Wiggins, 1993a, p.3) et ne devrait donc pas demander de ressources supplémentaires. Il associe ces résistances à une réticence aux changements: «*Genuine tests can be widely implemented if we can overcome inertia and fatalism about current forms of standardized testing.*» (Wiggins, 1989b, p.45) Qu’en disent les recherches ayant porté spécifiquement sur l’évaluation des coûts de la mise en place d’ÉSA et les comptes-rendus de praticiennes et praticiens ayant tenté de telles implantations dans leur classe ?

Peu de chercheurs ont travaillé spécifiquement sur la documentation des coûts financiers reliés à l'implantation d'ÉSA. Toutefois, à partir d'expériences menées dans différents États américains, certaines estimations ont été réalisées évaluant le coût d'implantation à large échelle d'une approche évaluative fondée sur l'évaluation des performances entre 2 et 4 fois (Hardy, 1996, p.107; Rothman, 1995, dans Hoepfl, 2000) voire entre 10 (Office of Technology Assessment, 1992, dans Hardy, 1996) et 60 fois (Office of Technology Assessment, 1992, dans Hardy, 1996) plus cher que l'approche centrée sur les tests standardisés.

Hardy (1996) identifie 3 étapes d'implantation d'une ÉSA - l'élaboration, l'administration et l'évaluation - ayant chacun leurs enjeux spécifiques en termes de ressources :

1. *L'élaboration* : créer un canevas d'ÉSA demande beaucoup de temps et d'énergie de la part d'un personnel qualifié, car il faut d'abord identifier les éléments de compétence et les objectifs d'apprentissage visés par l'activité, identifier des tâches permettant d'atteindre ceux-ci et des indicateurs adéquats pour en valider l'atteinte ou non, créer un cadre de réalisation adapté, rédiger des barèmes d'évaluation critériés pour chaque indicateur de performance, peaufiner le contexte de réalisation, créer les outils d'accompagnement et s'assurer de la viabilité globale de l'exercice (prétest et contrôle de la qualité). Selon Hoepfl (2000), les données actuellement disponibles concernant cette première étape d'élaboration d'une ÉSA indiquent qu'il en coûte entre 5 000 et 14 000\$ par épreuve créée. De plus, tous les chercheurs et chercheuses ayant étudié la question des coûts reliés la création et au contrôle de la qualité d'une ÉSA s'entendent sur un point : ils sont très difficiles à estimer et varient grandement selon le type de tâche exigée et de compétences impliquées (Hardy, 1996; Maeroff, 1991; Monk, 1996). Et le temps et les efforts consacrés pour cette étape sont peu compressibles, que le nombre d'étudiantes et d'étudiants auxquels il sera administré soit petit ou grand;

2. *L'administration* : la phase de la réalisation de l'ÉSA par les étudiantes et étudiants implique aussi plusieurs coûts. Tout d'abord, la mise en place d'un contexte authentique peut nécessiter l'utilisation de ressources matérielles ou humaines, ou encore la nécessité de se rendre dans un lieu hors de l'école afin d'augmenter l'authenticité de la tâche, ce qui peut engendrer des coûts. Ensuite, selon les principes de l'ÉSA, il faut prévoir un accompagnement pédagogique adéquat étant donné la nécessité d'une rétroaction fréquente et d'un soutien constant à l'apprentissage, ce qui peut signifier plus de ressources professionnelles adéquatement formées ou encore la limitation du nombre d'étudiantes et étudiants par groupe, ce qui a une incidence budgétaire importante (Hardy, 1996; Hoepfl, 2000; Khattri *et al.*, 1998; Monk, 1996);
3. *L'évaluation* : analyser l'ensemble de la production ou performance finale d'une ÉSA à partir de chaque indicateur afin de sanctionner l'atteinte ou non des objectifs d'apprentissage demande beaucoup plus de ressources que la correction d'un test standardisé. Tout d'abord, comme l'évaluation en situation authentique nécessite le recours au jugement professionnel, cette étape ne peut être faite de façon électronique ou par un personnel non qualifié : ainsi, le coût de la main-d'œuvre en mesure de réaliser cette évaluation, ou le coût de formation de personnel additionnel, sont considérables, d'autant plus que le temps nécessaire à la réalisation de cette étape est substantiel et difficilement compressible. En effet, comme une évaluation de chaque indicateur doit être réalisée à partir d'une production ou performance complexe, il faut prévoir beaucoup plus de temps que pour la validation d'une question à choix de réponse qui n'implique qu'une évaluation binaire très rapide (Hardy, 1996; Hoepfl, 2000; Kerka, 1993; Khattri *et al.*, 1998; Monk, 1996). L'impact sur les ressources d'une implantation d'ÉSA à large échelle est si grand que l'État du Kentucky a engagé une firme externe pour réaliser l'évaluation sommative finale lors d'un projet d'implantation d'une ÉSA mené au début des années 90 (Hoepfl, 2000). Ceci pose d'ailleurs la question de l'organisation du travail : si cette tâche n'est pas donnée en sous-traitance, il faudra prévoir des périodes où les enseignantes et enseignants pourront réaliser l'étape de

l'évaluation sommative dans les temps requis. Lors de projets-pilotes menés aux États-Unis et en Angleterre dans les années 1980 et 1990, le temps nécessaire à l'étape de l'évaluation sommative a été identifié comme le principal coût relié à l'implantation d'ÉSA, et ce pour deux raisons. La première découle de la nature du travail à évaluer : comme l'ÉSA exige une production ou performance complexe, le temps nécessaire à l'examen de celle-ci est plus important, en plus de la nécessité de s'adapter à l'unicité du résultat obtenu par chaque étudiante et étudiant (Hardy, 1996; Hoepfl, 2000; Maeroff, 1991; Popham, 1993). La seconde provient de l'ampleur de la rétroaction que prescrivent les principes de l'ÉSA : « *Diagnostic feedback on the quality of individual responses is seen by advocates of performance assessment as one of the method's most promising benefits. For large-scale assessments, however, the costs for providing this level of detail in scoring is likely to be prohibitive.* » (Hardy, p.114) Et en plus d'être substantiels, les coûts reliés à l'étape de l'évaluation sont récurrents et peu compressibles (Hardy, 1996; Maeroff, 1991; Monk, 1996; Popham, 1993).

On constate donc que les enjeux entourant la quantité, le type et la répartition des ressources nécessaires à la mise en place d'ÉSA sont non négligeables. Comme le synthétise Popham (1993) : « *in short, it takes money – and a lot of it.* » (p. 472)

Les tenants de l'approche répondront que l'adoption de l'ÉSA n'a pas pour objectif de réduire les coûts reliés à l'éducation, mais bien d'améliorer la qualité de l'apprentissage et de l'évaluation, qu'il s'agit en fait d'un investissement global, et non d'une simple dépense reliée à l'évaluation.

Yet many believe that the additional costs are a good investment. California Assessment Director Dale Carlson noted that 60% of the costs of performance assessment goes directly to teachers for their participation in scoring. He considers that money a worthy investment in staff development: "Teacher involvement brings you support, new ideas, and most of all, rejuvenation" (Hymes, 1991, p. 45). [...] The substantial additional costs are not considered as an assessment cost only, but as an

investment in the improvement of instruction and consequent student learning. Some argue that the alternative of continuing the current dependence on multiple-choice tests will ultimately be far more costly in terms of the constraints of that mode of assessment on school restructuring and on improved instruction for higher order thinking (Schulz, 1992). (Hardy, 1996, p. 116)

Dans un de ses textes les plus récents, Wiggins (2011), appuie cet argumentaire :

We must show that influential ‘‘monitoring’’ tests are so irrelevant (and even harmful) to genuine intellectual standards that their cost – to student learning and teacher professionalism – is too high, however financially efficient they may be as a means of gathering data. The inescapable dilemma presented by mass testing is that using authentic standards and tasks to judge intellectual ability is labor-intensive and time-consuming. Examiners must be trained, and multiple, contextual tests of the students must be conducted. Genuine Tests also make it more difficult to compare, rank, and sort because they rarely involve on simple, definitive test with an unambiguous result and a single residue number. Therefore, as long as tests are thought of only in terms of accountability, real reforms will be thwarted. (p.82)

Malgré tout, Khattri *et al.* (1998) rappellent que la question des ressources ne peut être esquivée, car «*the adequacy of the amount of time allowed for development, introduction, and institutionalization of assessment reform can have a dramatic impact on a state's ability to sustain its reform efforts.*» (p.74) Dans le contexte de l'enseignement supérieur au Québec, particulièrement dans les programmes préuniversitaires collégiaux et dans les baccalauréats où la tendance est à la réduction des ressources et à l'augmentation du ratio professeur-étudiant, cet enjeu questionne la faisabilité même de l'implantation de cette approche si aucun soutien institutionnel n'est offert aux enseignantes et enseignants et aux départements souhaitant adopter cette orientation.

Devant cette problématique des ressources limitées, particulièrement dans les contextes d'enseignement à un nombre élevé d'étudiantes et d'étudiants, certaines autrices et certains auteurs proposent des pistes de solutions.

Pour Montgomery (2002), la façon de rédiger les ÉSA peut faciliter le travail des étudiantes et étudiants, et conséquemment celui des enseignantes et enseignants.

The use of authentic tasks and rubrics that clearly state evaluation criteria for students can effectively break larger, lecture-oriented classes into smaller learning units. Rubrics can be used in peer review tasks as well as becoming a self-assessment tool. Authentic assessment is not meant to replace traditional testing, but rather to add to the multifaceted approach of contemporary assessment. (Montgomery, 2002, p. 36)

Hardy (1996) et Wiggins (1990, 2011) proposent, quant à eux, une approche centrée sur la coopération des étudiantes et étudiants et les activités de groupes :

Another approach to controlling materials costs may be to test collaborative groups, rather than individual students. The Georgia prototypes were developed for administration to groups of two to four students at considerable savings in costs per classroom.» (Hardy, 1996, p. 112)

Wiggins ajoute que la pratique de l'échantillonnage peut aussi être une solution : en évaluant sommativement une partie seulement de la production finale, on réduit le temps consacré à cette tâche. Toutefois, ces deux solutions créent des failles dans la capacité de certifier l'atteinte des objectifs d'apprentissage pour chaque étudiante et étudiant, et donc d'attester de sa capacité à poursuivre son cheminement académique à un niveau supérieur. Sur le plan des principes évaluatifs, cela va également à l'encontre du principe promu au niveau collégial quant à l'importance du caractère individuel des évaluations sommatives (Gouvernement du Québec, 2012; Hermann Guy, s.d.).

Finalement, mentionnons qu'au-delà de la quantité de ressources nécessaires, la mise en place d'ÉSA pose également des questions logistiques reliées à l'aménagement des structures et règles institutionnelles, notamment au niveau des politiques institutionnelles en matière d'évaluation et de l'organisation des calendriers scolaires. Nous voici donc au carrefour entre les enjeux de ressources et les contraintes institutionnelles.

4.3. Le contexte et les contraintes institutionnelles :

Pour Herrington et Herrington (2006), il y a deux principaux obstacles à l'adoption de l'ÉSA dans l'éducation postsecondaire : la première de nature institutionnelle et l'autre de nature pédagogique. Sur le plan institutionnel, ces auteurs constatent que dans certaines universités, les politiques concernant les modalités d'évaluation n'offrent pas aux enseignantes et enseignants la latitude nécessaire pour pouvoir implanter des ÉSA. Dans les dernières décennies, les mécanismes visant un contrôle de la qualité et une imputabilité plus grande ont amené certaines universités à exiger de la part des enseignantes et enseignants qu'ils se conforment à des normes très précises en matière d'évaluation - par exemple, des exigences quant à un nombre minimal d'évaluations et le respect d'un calendrier rigide ou la restriction quant au nombre d'étudiantes ou d'étudiants par équipe - dont certaines peuvent aller à l'encontre de l'approche par compétences et du paradigme constructiviste.

Despite much research to show that there are benefits to be derived from collaboration (Qin, Johnson, & Johnson, 1995), assessment strategies that take advantage of the use of group work can often be used only minimally because of assessment policies that limit the use of group work. In such cases, the focus is on each student learning—and being assessed—independently of the social context in which learning takes place. Of course, economic constraints may also compound the problem where university courses over rely on lectures that necessarily inhibit collaborative learning. (Herrington et Herrington, 2006, p. 148)

Ainsi, les procédures associées à des démarches de standardisation et de contrôle de l'enseignement, telle l'assurance-qualité – un courant qui s'implante d'ailleurs actuellement au Québec (Gouvernement du Québec, 2013) – peuvent amener des entraves à l'implantation d'ÉSA (Clarke *et al.*, 2000; Herrington et Herrington, 2006).

4.4. Les problèmes de nature pédagogique :

En plus de contraintes institutionnelles, Herrington et Herrington (2006) identifient certaines difficultés de nature pédagogique, reliées à l'élaboration d'une ÉSA et à de possibles interprétations erronées de la part des pédagogues et praticiennes et praticiens de ce qu'est l'évaluation en situation authentique.

Une première erreur recensée par ces autrices et ces auteurs est qu'un scénario conçu par l'enseignante ou l'enseignant peut introduire involontairement des biais dans l'authenticité de la situation en ajoutant des éléments de contexte ou d'évaluation qui détournent ou distraient les étudiantes et étudiants des réelles connaissances et habiletés centrales reliées à la tâche demandée. Dans ces cas, les éléments du scénario prennent le dessus sur les exigences extrascolaires normalement associées à une telle tâche et les attentes d'ordre secondaire prennent le dessus sur celles de premier ordre (Cumming et Maxwell, 1999; Herrington et Herrington, 2006; Newmann et Archbald, 1992; Wiggins, 1993*a*, 1993*b*), ce qui rejoint les enjeux de validité énoncés précédemment.

Un autre accroc de nature pédagogique identifié par Cumming et Maxwell (1999) est le camouflage, c'est-à-dire la fausse contextualisation dissimulant une tâche non authentique. Non seulement ce n'est pas utile, mais cela peut ajouter de la confusion chez les étudiantes et les étudiants :

Many interpretations of authenticity verge on labelling and exultation by association, rather than capacity to reduce rather than improve the quality of teaching and assessment. A return to fundamental principles is urgently

needed. Simplistic implementations of authentic assessment through camouflage, and complex but superficial implementations through simulation, miss the mark and do not improve educational practice. (Cumming et Maxwell, 1999, p. 192-193)

Finalement, Messick met en garde contre deux problèmes qui peuvent survenir dans l'élaboration d'une ÉSA, soient

- a) la sous-représentation de certains éléments de la compétence, qui survient lorsque *«the assessment is too narrowly focused, failing to include important dimensions of the knowledge or skill it aims to assess.»* (Messick, 1996, p. 5);
- b) l'erreur de pertinence, qui survient lorsque la tâche appelle des réponses qui ne sont pas pertinentes pour mesurer les connaissances et habiletés souhaitées. (Messick, 1996, p. 5).

Pour pallier ces risques, Crocker (1997, dans Hoepfl, 2000) propose deux critères de validité d'une ÉSA, qui rejoignent d'ailleurs les préoccupations de Messick. Pour Crocker, une ÉSA doit pour être valide : a) être composée de tâches et de critères d'évaluation pertinents en regard des connaissances reconnues dans le domaine d'expertise concerné, et b) couvrir de façon équilibrée et représentative l'étendue des dimensions de la réalité étudiée.

On constate donc que les problèmes didactiques et pédagogiques associés à la création d'ÉSA ne sont jamais bien loin des questions plus théoriques mentionnées jusqu'à maintenant, notamment le problème de l'absence d'uniformité et de consensus à propos de la définition de l'authenticité d'une évaluation et des problèmes de validité. Car comme le souligne Terwilliger (1997) à propos du modèle d'ÉSA élaboré par Wiggins: *« it is not entirely clear what is meant by the phrase, "designed to be truly representative of performance in the field." What exactly is the "field" in this example? Who decides what is "truly representative" of the field? This raises questions regarding test validity.»* (p. 26)

Finalement, un problème de nature pédagogique provient du manque d'expertise des étudiantes et des étudiants en matière d'ÉSA. Montgomery (2002) mentionne que certains praticiennes et praticiens ont relevé des difficultés dans la réalisation de l'ÉSA en raison du peu d'expérience des étudiantes et des étudiants avec ce type d'évaluation. La nécessité de rédiger des réponses ouvertes et inédites, la grande latitude dont disposent les étudiantes et étudiants, le caractère appliqué de la tâche et le niveau d'habileté intellectuelle impliqué semblent causer de l'insécurité et un grand besoin d'accompagnement, faute d'avoir été souvent exposés à cette formule (Liftig, Liftig et Eaker 1992, dans Montgomery (2002). Parmi les pistes de solutions émises afin de faire face à cet enjeu, on retrouve l'augmentation de l'exposition des étudiantes et étudiants à des ÉSA et une augmentation des ressources en phase d'implantation.

L'examen des problèmes de mise en œuvre des ÉSA nous révèle que l'intégration de ce concept dans la pratique n'est pas encore totalement maîtrisée et que plusieurs difficultés doivent encore être résolues. On remarque aussi que certains promoteurs de l'ÉSA - dont Wiggins - tentent plutôt de minimiser ces difficultés plutôt que de les aborder de front, ce qui peut être dangereux, comme le souligne (Ewing, 1998) :

Many writers worry that the current popularity of alternative assessment may lead to its eventual downfall. They fear that educators are so enthusiastic about the potential of the newer forms of assessment that they are ignoring some of the possible pitfalls. Bracey (1993) fears that "unless the proponents for 'authentic' assessments are as 'tough-minded' in developing them as they have been in critiquing traditional practice, the new tests may not be professionally credible, publicly acceptable, or legally defensible" (p. 9). Worthen (1993) warns the alternative assessment community to attend to the essentials: conceptual clarity, mechanisms for self-criticism, support from well-informed educators, technical quality and truthfulness, standardization of assessment judgments, ability to assess complex thinking skills, and acceptability to stakeholders. For Maeroff (1991) and Baker (1993), a key issue is the appropriate use of alternative assessments. Both charge that educators are blindly assuming that alternative assessments are appropriate in all cases, which may not be true. (p. 14). (Ewing, 1998, p.11)

Ceci ne fait que renforcer la pertinence de la démarche de cet essai, qui est de documenter les conditions favorisant la mise en œuvre d'évaluation en situation authentique à travers les difficultés rencontrées des pistes de solution identifiées par des enseignantes et des enseignants du DEC en sciences humaines qui ont tenté cette aventure.

Pour conclure cette section sur le cadre de référence, rappelons que c'est le concept d'évaluation en situation authentique qui est au cœur de la démarche, de la collecte de données et de l'analyse. Pour bien situer ce concept, les origines et l'évolution ont été présentées. Un portrait de la littérature anglophone et francophone disponible sur le sujet a été réalisé, ainsi qu'une analyse critique de celle-ci. Le choix de travailler principalement à partir des modèles provenant d'auteurs anglophones, plus spécifiquement le modèle de Wiggins et celui d'Archbald et Newmann, a été justifié, et une synthèse de ces deux modèles, ainsi qu'une réflexion sur leurs apports respectifs et sur leur degré de compatibilité avec la présente démarche, ont été effectuées. Finalement, les difficultés, ainsi que les conditions et actions favorables pour les contourner, reliées à la mise en œuvre d'ÉSA telles que documentées par les spécialistes de ce concept et par des praticiennes et praticiens ayant expérimenté des implantations d'ÉSA dans leur milieu académique, ont été recensées. Le tableau 1 situé à l'annexe E présente d'ailleurs une synthèse de ce dernier point.

5. LES OBJECTIFS SÉCIFIQUES DE L'ESSAI

Considérant l'état des connaissances et des questionnements sur l'évaluation en situation authentique, l'objectif de recherche principal suivant a été retenu : documenter les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans des cours de première année du programme d'études préuniversitaires en sciences humaines au collégial.

À cet objectif général se greffent trois objectifs spécifiques :

1. Identifier des conditions et actions perçues comme favorables à la mise en œuvre de pratiques évaluatives s'appuyant sur les principes d'une évaluation en situation authentique par une enseignante ou un enseignant ayant expérimenté la mise en œuvre d'une telle évaluation dans le cadre d'un cours de première année du programme d'études préuniversitaires en sciences humaines;
2. Identifier les enjeux et difficultés rencontrées par une enseignante ou un enseignant ayant expérimenté la mise en œuvre d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans le cadre d'un cours de première année du programme d'études préuniversitaires en sciences humaines;
3. Analyser les enjeux et difficultés rencontrées et les conditions et actions perçues comme favorables identifiés par une enseignante ou un enseignant durant l'implantation d'une évaluation terminale de cours en situation authentique dans le cadre d'un cours de première année du programme d'étude préuniversitaire en sciences humaines.

TROISIÈME CHAPITRE LA MÉTHODOLOGIE

Ce chapitre présente la démarche méthodologique mise en œuvre afin de documenter les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans des cours de première année du programme d'études préuniversitaires en sciences humaines au collégial. Tout d'abord, le type d'essai retenu est présenté, suivi des considérations concernant la posture épistémologique adoptée dans la démarche de recherche. S'en suivent le type de recherche et l'approche méthodologique, suivis des aspects éthiques associés à la démarche et à la collecte de données, et de la présentation des moyens pris pour assurer la scientificité.

1. TYPE D'ESSAI

Cet essai prend la forme d'une étude de cas et porte sur des initiatives d'implantation d'épreuves terminales de cours en situation authentique (ÉTCSA) menées individuellement par une un professeur qui a librement choisi de le faire dans le cadre de cours de première année du DEC en sciences humaines. Selon Roy (2010), l'étude de cas est «une approche de recherche empirique qui consiste à enquêter sur un phénomène, un événement, un groupe ou un ensemble d'individus, sélectionné de façon non aléatoire, afin d'en tirer une description précise et une interprétation qui dépasse ses bornes.» (p. 206–207) Pour Gauthier (2010),

cette approche de recherche se caractérise à la fois par le nombre restreint de situations analysées, la profondeur de l'analyse et l'importance accordée à une démarche inductive, qui alimentera une phase de développement de théories ou de modèles (Gauthier, 2010, p. 173).

Paillé adapte la définition de l'étude de cas au contexte des sciences de l'éducation en indiquant que

[l]'étude de cas est une forme d'enquête (socio-pédagogique) ou d'investigation (psycho-pédagogique) portant sur un petit nombre de cas, parfois sur un seul. Le cas peut être représenté par une personne — par exemple une étude de cas clinique en orthopédagogie — ou par une classe (étude d'un cas de pédagogie par projet), une direction d'école (étude d'une forme de gestion participative), etc. Cette forme de recherche est particulièrement bien adaptée à la maîtrise puisqu'elle permet d'étudier un phénomène plus en profondeur que ne le permet l'enquête à grande échelle ou la recherche expérimentale (Dufour, Fortin et Hamel, 1991). (Paillé, 2007, p. 146)

Dans le cadre de cet essai, nous analyserons une expérience d'implantation d'ÉTCSA par un professeur enseignant au moins un cours de première année du programme préuniversitaire en sciences humaines afin d'en faire ressortir les caractéristiques et d'identifier des pistes de facteurs favorisant et défavorisant la réussite de ce type d'initiative.

Le choix de l'étude de cas comme type d'essai est notamment motivé par des considérations d'ordres contextuel, logistique et méthodologique. Tout d'abord, l'impossibilité d'étudier la problématique retenue à grande échelle en raison du petit nombre de sujets disponibles et des ressources limitées. Ensuite, le peu de documentation sur ce thème dans le contexte collégial et québécois, d'où le caractère exploratoire de cette recherche. Cette absence de documentation, malgré l'existence d'une réelle problématique, rend ardu le recours à un essai de type expérimental en raison de la difficulté d'émettre des hypothèses significatives et précises à vérifier : il est donc préférable d'opter pour une méthode propice à une exploration de la problématique, ce qui est possible avec l'étude de cas (Roy, 2010). Comme le souligne Fortin (2010), «[l]'étude de cas est appropriée quand on dispose de peu de données sur la personne, l'événement ou le phénomène considéré (Yin, 2002)». (p. 295)

L'importance d'étudier le phénomène dans son contexte est la seconde raison du choix de l'étude de cas comme type d'essai. Étant donné que nous nous intéressons à l'implantation d'ÉTCSA dans un cadre spécifique – les cours prévus à la grille lors de la première année du programme de sciences humaines au collégial⁴ – il est très important d'étudier ce cas dans son contexte propre. De plus, comme le cas à l'étude est une ÉTCSA conçue et implantée dans le cadre d'une initiative individuelle, selon un modèle singulier, il est fondamental de bien en saisir les détails en le situant constamment dans son contexte de réalisation. Or,

[c]ette méthode de recherche [l'étude de cas] est particulièrement appropriée pour des questions pratiques où l'expérience des acteurs est importante et son contexte essentiel à connaître (Benbasat, 1984; Benbasat et coll., 1983; Bonoma, 1983; Roethlisberger, 1977. Yin (1981a, p. 98) écrit : "The need to use case studies arises whenever : an empirical inquiry must examine a contemporary phenomenon in its real-life context are not clearly evident" (Gagnon, 2012, p. 15)

Finalement, la difficulté de recruter un grand nombre de sujets a confirmé la pertinence de l'étude de cas pour la présente démarche de recherche, car, comme le souligne Fortin (2010), cette méthodologie permet de faire de cette contrainte une force :

Sa valeur réelle réside dans le caractère poussé de l'analyse, les multiples observations auxquelles elle donne lieu et les comportements types qu'elle permet d'isoler. Au nombre des avantages de l'étude de cas, on compte l'information détaillée que l'on obtient sur un phénomène nouveau, les idées qu'elle permet de dégager, l'établissement de relations entre les variables et la possibilité de conduire à la formation d'hypothèse. L'étude de cas peut également devenir une source d'information valable pour la pratique fondée sur des données probantes du fait qu'elle fournit des aspects importants d'un point de vue clinique ou social. Ses limites viennent du fait que les résultats ne peuvent être généralisés à d'autres

⁴ Étant donné la variation entre les cégeps au niveau de la grille de cours, il n'est pas possible de faire l'énumération complète des cours et matières concernés : le critère utilisé pour la sélection des participants sera le fait d'enseigner un cours prévu pour la première année du programme de sciences humaines dans la grille de cours de leur institution.

populations ou situations et que les données peuvent être insuffisantes ou difficilement comparables entre elles. (p. 296)

Pour toutes ces raisons, l'étude de cas se présente comme le type d'essai le plus compatible avec le type de problématique en cause et le contexte dans lequel elle se manifeste.

Quant au type d'étude de cas qui sera privilégié, comme le but de la démarche est de faire ressortir des pistes de réflexion sur les facteurs favorisant et défavorisant l'implantation d'ÉTCSA (qui pourront éventuellement devenir des hypothèses pour une recherche subséquente) à partir du cas étudié, l'étude de cas unique et instrumentale a été privilégiée; en effet, la démarche de recherche se centre sur une problématique, est subordonnée à un intérêt externe au cas étudié et fait appel à un cas pour l'analyse de la démarche (Fortin, 2010; Karsenti et Savoie-Zajc, 2011). Finalement, selon la typologie de Merriam, cette démarche se catégorise comme traitant d'un problème d'action, lesquels «proviennent d'un besoin de solution de rechange pour résoudre un problème concret» (Karsenti et Savoie-Zajc, 2011, p. 241), dans le présent cas, la faible perméabilité à l'évaluation dans une approche par compétences rapportée par la CÉEC et plus précisément la lente implantation des ÉTCSA dans les cours de sciences humaines.

2. LA POSTURE ÉPISTÉMOLOGIQUE

L'approche méthodologique étant fortement influencée par la posture épistémologique du chercheur, il importe d'aborder ce point. La présente démarche de recherche se situe dans le paradigme interprétatif, avec quelques influences de l'approche critique. Voici en détail le positionnement pour chacun des thèmes qui définissent les spécificités des paradigmes épistémologiques selon Karsenti et Savoie-Zajc (2011).

2.1 La vision de la réalité

Pour ce qui est de la vision de la réalité, cet essai se campe dans une approche interprétative, avec influence de l'approche critique. Ainsi, le postulat du paradigme interprétatif selon lequel «[l]a réalité est construite par les acteurs d'une situation; elle est globale, car c'est la dynamique du phénomène étudié que le chercheur veut comprendre» (Karsenti et Savoie-Zajc, 2011, p. 115) est privilégié dans cette démarche; en ce sens, ses assises épistémologiques sont principalement interprétatives. Toutefois, la perception de l'approche critique selon laquelle certaines influences extérieures à la seule compréhension élaborée par les acteurs viennent influencer le sens que ceux-ci donnent à leur façon d'appréhender cette réalité et à leurs actions est à notre sens un apport important de cette posture épistémologique, qui de plus est compatible avec l'essence de la posture interprétative. Ainsi, dans l'approche du phénomène étudié, l'influence des structures et des cadres existants dans la construction interprétative des acteurs d'une réalité sera considérée, ainsi que de l'influence des rapports de force dans celle-ci. Il importe toutefois de rappeler que l'approche adoptée dans la présente recherche s'éloigne du déterminisme parfois associé à la posture critique dans sa forme radicale, car, bien que les structures et cadres influencent la perception des acteurs, ceux-ci ont une grande liberté dans leur façon de donner du sens, tout particulièrement dans le contexte collégial où les enseignantes et enseignants disposent d'une liberté académique non négligeable.

2.2 La nature du savoir et la place du chercheur

Conformément à la posture interprétative, la présente démarche s'inscrit dans une conception du savoir intimement liée au contexte de sa production et transférable à d'autres contextes. La production du savoir ne peut être parfaitement neutre, car le chercheur ne peut se dégager totalement de ses valeurs : ainsi, «[d]ans cette démarche, le chercheur et les participants à la recherche ne sont pas neutres : leurs schèmes personnels et théoriques, leurs valeurs influent sur leur conduite et le chercheur tente de produire un savoir objectivé, c'est-à-dire validé par les participants à la recherche.»

(p.128) Étant donné ce constat, celui-ci «révèle ses approches face à sa recherche [...] cherche à objectiver ses données.» (Karsenti et Savoie-Zajc, 2011, p. 115), et c'est en partie l'objectif de la présente section.

Pour le paradigme interprétatif, les sujets prenant part à l'étude sont la clé de voute de la qualité du savoir obtenu et sont souvent inclus dans le processus de validation des analyses. «Une recherche est d'ailleurs jugée valable et crédible si les gens qui y ont contribué s'y reconnaissent.» (Karsenti et Savoie-Zajc, 2011, p. 125) C'est ici que notre adhésion au paradigme interprétatif s'affaiblit et que nous sentons le besoin d'une nouvelle hybridation avec l'approche critique. Bien que nous adhérons entièrement à cette idée de l'informateur spécialiste de sa propre réalité lors du processus de collecte et de présentation des données, où il est effectivement primordial que les participants perçoivent le portrait réalisé par le chercheur comme représentatif de leur réalité et que les participants soient effectivement source de validité pour cette phase, l'interprétation des données recueillies par le chercheur ne doit pas avoir cet objectif, sous peine de devenir complaisante vis-à-vis de la perception des participants à l'étude.

Pour ce qui est de l'évaluation de l'interprétation des résultats, les participantes et les participants ne peuvent être la source principale de validation des données : la contribution du chercheur doit aller plus loin et dégager - grâce à l'analyse et aux liens avec les recherches et théories extérieures à la collecte de données qu'il a réalisées - des éléments qui peuvent être en dissonance avec la perception des participants de la recherche. C'est pour nous la seule façon de faire ressortir des influences structurelles parfois taboues ou non identifiées par les acteurs. Donnons comme exemple les études sur le racisme ou le sexisme : on ampute considérablement la portée des recherches si on ne permet pas aux chercheurs d'analyser les données fournies par les participantes et les participants d'une recherche en s'autorisant à effectuer des interprétations de propos avec lesquels les participants pourraient être en désaccords (par exemple, le chercheur pourrait classer certains propos comme racistes alors que le participant nierait féroceement cette étiquette). Ainsi, un apport important

de l'approche critique est de permettre la mise à jour de certaines influences provenant des structures et normes extérieures à la situation et qui influencent notre compréhension du réel. C'est d'ailleurs ce qui donne son pouvoir émancipatoire à la recherche : si le sujet est l'expert de la description et de l'interprétation de son vécu, la chercheuse ou le chercheur est le mieux placé pour produire une analyse objectivée (bien que jamais parfaitement neutre) à partir de cette information fournie par le sujet.

Bref, la posture interprétative sur la nature du savoir et ses sources de validation permet d'évaluer la qualité de la démarche de collecte et de la préparation des données recueillies dans la recherche, mais pour ce qui est de l'interprétation des données, la distance prise par le paradigme critique nous semble essentielle et au cœur du rôle du chercheur dans la production de la connaissance.

2.3 Finalité de la recherche

Dans un premier temps, la finalité de la recherche doit être de «comprendre la dynamique du phénomène grâce à l'accès privilégié du chercheur à l'expérience de l'autre» (Karsenti et Savoie-Zajc, 2001, p. 115), car si on adopte dès le départ une posture visant une finalité critique avant même d'avoir eu pour objectif de vouloir comprendre le phénomène étudié, il y a un grand danger d'être plus vulnérable aux biais de recherche liés aux préconceptions. Mais une fois cette première finalité atteinte, la «seconde phase» de la finalité de la recherche est de permettre à l'enseignante-praticienne ou l'enseignant-praticien de mettre en perspective ses perceptions et la compréhension qu'il a de la situation étudiée afin d'en distinguer les influences des structures dominantes et de pouvoir et s'il le considère juste, prendre une distance vis-à-vis d'elles et éventuellement transformer les aspects qu'il perçoit comme critiquables. À notre avis, pour ce qui est de la finalité de la recherche, chaque posture a son moment approprié, ce qui explique que nous optons pour un cadre épistémologique hybride. Le chercheur trop centré sur le changement durant la collecte et le début de l'analyse risque de ne pas suffisamment objectiver ses données; à

l'inverse, celui qui n'ouvre pas ses horizons aux impacts de ses découvertes sur les transformations du monde souhaitables ne permet pas une pleine optimisation des retombées de ses recherches pour la communauté. Mais il est important de le redire : chaque posture en son temps.

3. TYPE DE RECHERCHE ET APPROCHE MÉTHODOLOGIQUE

La présente recherche est de nature non expérimentale à visée exploratoire et idiographique.

Par ailleurs, la démarche reposera sur une question et un objectif de recherche exploratoire menant à une structure de preuve inductive. (Gauthier, 2010) synthétise ainsi les caractéristiques d'une recherche à visée exploratoire :

Les questions de recherche exploratoire visent des thèmes qui ont été peu analysés et dont le chercheur n'est pas en mesure d'établir un portrait à partir des connaissances existantes. [...] L'objectif de recherche, dans le cas des questions exploratoires, est de nature inductive : la situation existante peut-elle nous apprendre quelque chose que l'on puisse formuler ensuite au moyen d'un modèle temporaire de représentation de la réalité ? Pour aborder les questions exploratoires, on privilégie une approche qui permet de s'imprégner de l'essence d'une situation, d'en capter la complexité et d'en interpréter le sens. L'approche exploratoire par excellence est l'étude de cas. (p. 171-172)

Ajoutons également que cette recherche présente une visée idiographique plutôt que nomothétique : «[l]a première cherche à comprendre un phénomène dans son contexte, tandis que la seconde s'efforce d'établir des lois générales (Franz et Robey, 1984; Weick, 1984).» (Gagnon, 2012, p. 14) Ceci est parfaitement cohérent avec une approche qualitative exploratoire de type étude de cas.

3.1 Les critères de sélection

La présente étude porte sur l'expérience de mise en œuvre d'ÉTCSA dans le cadre de cours de première année du programme préuniversitaire en sciences humaines dans les cégeps du Québec. La population cible est donc composée des protagonistes de ces expériences : les enseignantes et les enseignants ayant initié de façon libre et individuelle de telles expériences.

La recherche d'un ou une volontaire s'est donc faite à partir des critères suivants : le participant a expérimenté l'implantation d'au moins une évaluation terminale de cours conforme aux principes de l'évaluation en situation authentique

- a) élaborée et mise en œuvre de sur une base volontaire et individuelle;
- b) entre le 1^{er} janvier 2015 et le 31 décembre 2017 ;
- c) dans le cadre d'un cours du programme préuniversitaire en sc. humaines prévu à la 1^{re} année du programme d'étude dans la grille de cours du cégep de provenance
- d) dans un cégep public;
- e) dans des classes où le nombre d'étudiantes et d'étudiants inscrits au premier jour de classe était d'au moins 30 étudiants;
- f) dans le cadre de cours qui ne s'inscrivaient pas dans un projet pédagogique spécial pour lequel les étudiantes et les étudiants avaient fait l'objet d'une sélection (par exemple, un cours où les étudiantes et étudiants sont sélectionnés pour la participation à un stage de mobilité à l'étranger).

Les expériences d'implantation collective d'ÉTCSA sont exclues, afin d'inclure uniquement des expériences menées sans contraintes, c'est-à-dire de manière libre, volontaire et individuelle, avec optimisme envers le principe d'évaluation en situation authentique, condition nécessaire pour repérer les facteurs facilitants ou entravants qui résultent du contexte général de l'enseignement de cours de 1^{re} année du programme d'études préuniversitaires en sciences humaines (condition

d'enseignement commune à tous, type de population étudiante, etc.) ou de difficultés inhérentes au principe même de ce type d'évaluation. Or, sélectionner une participante ou un participant ayant expérimenté l'implantation collective d'une ÉTCSA pourrait brouiller les résultats, et conséquemment l'analyse, en introduisant des influences propres à un contexte collectif (par exemple, la participation d'enseignantes ou d'enseignants ayant subi des pressions - voire ayant été contraints - de participer en raison de décisions départementales, ou encore le fait que lors d'une implantation collective chaque enseignante ou enseignant a moins de latitude individuelle pour la prise de décisions), ce qui s'éloigne de l'objet de recherche premier de cette démarche.

Pour les mêmes raisons, nous avons choisi d'étudier seulement une expérience dont l'enseignante ou l'enseignant participant a été le seul ou un des principaux concepteurs de l'ÉTCSA et avait le plein contrôle des paramètres reliés au déroulement de l'évaluation, afin d'éliminer des influences parasites reliées au manque d'emprise d'une professeure ou d'un professeur implantant un protocole conçu ou dirigé par une autre enseignante ou un autre enseignant.

Autre élément visant à réduire les variables parasites : le participant ou la participante devait enseigner dans un cégep public. Les différences dans le financement et les réalités des institutions collégiales publiques et privées a motivé ce choix : comme les budgets mis à la disposition des enseignantes et enseignants et l'organisation du travail de celles-ci et ceux-ci peuvent influencer grandement la capacité et la forme des activités pédagogiques en situation authentique et qu'il nous est impossible de disposer des détails qui permettraient de mettre en perspective le cas étudié en tenant compte de cette réalité, nous avons ciblé les enseignantes et enseignants des institutions publiques seulement, lesquelles sont plus nombreuses que leurs vis-à-vis privées.

Sur le plan des limites spatio-temporelles, la personne recrutée devait avoir réalisé l'expérience discutée lors d'au moins une session entre le 1^{er} janvier 2015 et le

31 décembre 2017. Cette limite vise en premier lieu à respecter la capacité de rappel du participant en ne traitant pas d'expériences trop éloignées dans le temps.

Finalement, le choix du cas étudié a été réalisé selon l'approche du choix raisonné,

qui consiste à sélectionner certaines personnes en fonction de caractéristiques typiques de la population à l'étude. [...] À partir des connaissances qu'il a de la population, le chercheur exerce son jugement sur le choix des personnes aptes à fournir l'information liée au but de l'étude. (Fortin, 2010, p. 235).

Toujours selon Fortin, «[l']échantillonnage par choix raisonné est la stratégie privilégiée dans l'étude de cas, considérant que le choix des participants et des sites est fait en fonction de la compréhension du problème de recherche et du phénomène à l'étude (Creswell, 2003)» (p. 281)

Bien qu'en recherche qualitative, un échantillonnage optimal est défini par l'atteinte du seuil de saturation (Fortin, 2010), l'ampleur de l'échantillon a été prédéterminée et établie à une enseignante ou un enseignant en raison du temps et des ressources disponibles pour la réalisation de la collecte de données ainsi que du caractère exploratoire de la visée de recherche. L'objectif de cet essai n'est pas d'atteindre la saturation des données, mais d'opter pour une approche exploratoire réaliste dans le cadre imparti. De plus, la méthodologie de l'étude de cas autorise le choix d'un nombre de participants fixe. Selon Roy (2010) : «[un] cas peut être sélectionné pour son caractère révélateur, son potentiel de découverte, ou pour l'opportunité qu'il présente d'étudier un phénomène en temps réel» (p. 215)

3.2 Le participant

Après des démarches de recrutement auprès de plusieurs enseignantes et enseignants travaillant dans différents cégeps, la personne retenue pour participer à notre étude est un professeur qui enseigne depuis plus de 10 ans une discipline du programme de sc. humaine au collégial, notamment auprès des étudiantes et des étudiants de première année de ce programme : au cours des deux dernières années, il estime travailler auprès de cette population dans une proportion de 25% à 49% de sa tâche. Il a un statut de professeur permanent dans un cégep de taille moyenne et a enseigné à temps plein durant au moins les deux dernières années. Afin de faciliter la présentation des résultats, tout en assurant la confidentialité des résultats, le participant sera désigné par le prénom fictif «Julien».

Au niveau de son profil professionnel, il a effectué des activités créditées de perfectionnement en pédagogie, certaines dans le cadre d'un programme de pédagogie régulier, ouvert à tous, dans une université québécoise, et d'autres dans le programme de formation continue offert via son employeur. Il a également suivi des formations non créditées en pédagogie, notamment dans le cadre de journées pédagogiques organisées par le cégep dans lequel il enseigne et en participant à des colloques. Lors de ces activités non créditées, il s'est intéressé à une variété de sujets d'ordre pédagogique et didactique.

Interrogé sur son niveau de connaissance du concept d'évaluation en situation authentique (Q11)⁵, il a déclaré avoir un excellent niveau de connaissance de ce concept, soit le plus haut niveau proposé parmi les modalités de réponse à cette question. Concernant le cadre dans lequel il a acquis ses connaissances sur ce concept,

⁵ La lettre Q désigne le questionnaire rempli par le participant, et 11 est le numéro de la question posée au participant pour obtenir cette information. Ce questionnaire est présenté en annexe B.

il a déclaré qu'elles ont été uniquement acquises dans le cadre d'un cours ou d'une formation en pédagogie (Q12)⁶.

Sur le plan de son expérience en matière d'ÉTCSA, Julien déclare avoir expérimenté cette approche auprès de 6 à 9 groupes d'étudiantes et d'étudiants de première année du programme de sciences humaines au cours des 2 dernières années, dans le cadre d'un cours de ce programme (Q9). Interrogé par questionnaire sur son expérience globale de mise en œuvre d'une ÉTCSA auprès d'étudiantes et d'étudiants de première année du programme de sciences humaines au cours des deux dernières années, il s'est déclaré globalement «assez satisfait» de ces expériences (Q14).

Au niveau des expériences d'implantation d'ÉTCSA auprès d'étudiantes et d'étudiants de première année du programme de sciences humaines réalisées par le Julien, elles ont toutes été réalisées à partir d'un même concept d'évaluation en situation authentique, qu'il a peaufiné et amélioré avec le temps. Avant d'aller plus en détail dans la présentation du participant et du cas à l'étude, il importe de faire une mise en garde : il n'est pas possible de faire une présentation détaillée du contexte de l'ÉTCSA qui a fait l'objet de la collecte de données, car une telle présentation permettrait facilement l'identification du participant par recoupement des informations, notamment du fait que son projet est disponible librement à la communauté collégiale sur un site de diffusion de matériel pédagogique. Il a donc fallu présenter certaines informations dans des catégories plus larges, et se restreindre aux critères d'admissibilité à l'étude, et omettre certains détails qui auraient permis d'identifier le participant par recoupement d'information, afin d'assurer le respect des normes éthiques associées à la confidentialité dans le cadre d'un essai de maîtrise, malgré la perte de précisions inhérente à ce choix. Les informations nominatives complètes

⁶ Pour répondre à cette question, le participant pouvait identifier jusqu'à quatre sources d'informations, soient l'étude dans le cadre d'un cours ou d'une formation (créditée ou non) en pédagogie, la lecture autonome à propos de ce concept, l'information provenant de collègue et l'information provenant d'un conseiller ou d'une conseillère pédagogique.

seront conservées selon les modalités prévues lors de l'émission du certificat d'approbation éthique autorisant la collecte de données émis par l'institution de provenance du participant.

3.3 Méthode de collecte des données

Trois méthodes de collecte de données ont été utilisées : un court questionnaire (présenté en annexe B), une entrevue semi-dirigée (dont le schéma est présenté en annexe C) et la collecte d'artefact (dont la grille de collecte est présentée en annexe D).

La collecte de données a été réalisée le 29 novembre 2017 dans les bureaux de l'établissement où enseigne Julien. Le déroulement a été le suivant :

1. Présentation de la chercheuse et de la recherche dans son ensemble;
2. Lecture du formulaire de consentement et précision concernant la nature de la recherche et la nature de la participation du répondant;
3. Validation de la grille d'évaluation de l'admissibilité à la recherche;
4. Présentation du questionnaire, que le participant complète ensuite;
5. Présentation et réalisation de l'entrevue;
6. Collecte de la documentation liée aux évaluations en situation authentique utilisée par le participant et inscription de ceux-ci dans la grille de collecte des documents.

3.3.1. L'entrevue semi-dirigée

L'entrevue semi-dirigée est l'élément central de la collecte de donnée, car cette méthode de collecte permet d'accéder à un savoir expérientiel en profondeur et contextualisé, en rendant l'univers de l'autre explicite, en permettant d'accéder à sa compréhension, en organisant et structurant la pensée de la personne interviewée et en jouant une fonction émancipatrice pour l'interlocuteur par l'approfondissement de son

vécu grâce aux réflexions que l'échange suscite (Savoie-Zajc, 2010). Cette méthode de collecte des données est donc à la fois appropriée au type d'informations recherchées – un savoir expérientiel individuel - et au type d'essai réalisé - l'entrevue est une des méthodes privilégiées pour réaliser des études de cas (Fortin, 2010; Gagnon, 2012; Roy, 2010). Son caractère semi-dirigé permet de donner à l'entrevue une certaine structure tout en laissant la flexibilité nécessaire pour adapter les questions au contexte et aux propos des participants. De plus, dans sa forme constructiviste, l'entrevue «[met] l'accent sur l'aspect construit de l'entrevue alors que chercheur et participants construisent un "texte", une "narration" constitués d'histoires, empreintes de connotations personnelles, interpersonnelles, sociales et culturelles des individus en présence.» (Savoie-Zajc, 2010, p.339) Finalement, cette méthode est compatible avec notre approche épistémologique, car «une dynamique de coconstruction de sens s'établit donc entre les interlocuteurs chercheur et participants. Les uns apprenant des autres et stimule l'émergence d'un nouveau discours et d'une nouvelle compréhension, à propos du phénomène étudié.» (Savoie-Zajc, 2010, p. 339)

Sur le plan de la conception du schéma d'entrevue, il s'agit d'un schéma construit expressément pour la réalisation de cette étude, divisé en trois parties. La première partie traite des expériences du participant en matière d'implantation d'évaluation terminale de cours en situation authentique. La seconde partie aborde le vécu en matière d'implantation d'une évaluation terminale de cours en situation authentique et a été structurée, non pas à partir des catégories de problèmes identifiés dans le cadre théorique et dans la littérature (voir Tableau 1 - Synthèse des problèmes reliés à la conception et la mise en œuvre d'une ÉSA recensés dans la littérature, Annexe E), mais plutôt à partir d'une structure reflétant le déroulement chronologique du processus de conception et de mise en œuvre de l'ÉTCSA par le participant. À la suite des contacts préliminaires avec Julien, au cours desquels il a brièvement exposé son cheminement, c'est le modèle de déroulement d'une ÉSA en trois phases – conception, administration et évaluation - de Hardy (1996) qui a été retenu pour structurer cette partie de l'entrevue, afin de faciliter le témoignage du participant, en se

collant à sa réalité et au déroulement naturel de son expérience. Pour chacune de ces phases, une attention particulière a été portée à documenter les enjeux et difficultés rencontrés dans le processus de mise en œuvre d'évaluations terminales de cours en situation authentique et les conditions et actions perçues comme favorables par le participant lors d'une telle démarche.

La troisième partie du schéma d'entrevue aborde les thèmes présents dans la littérature concernant les obstacles et les pistes de solution propres associés aux évaluations en situation authentique. Cette partie a été structurée à partir des catégories de problème recensées dans la littérature (présenté dans le Tableau 1 - Synthèse des problèmes reliés à la conception et la mise en œuvre d'une ÉSA recensés dans la littérature, Annexe E). Afin de ne pas prolonger indument l'entrevue, cette partie a été réalisée sous la forme de relance auprès du participant autour de ces thèmes. Ces relances n'étaient effectuées que si le thème en question n'avait pas été spontanément abordé par le participant, ou s'il n'avait pas été suffisamment approfondi dans les parties précédentes de l'entrevue.

En guise de conclusion à l'entrevue, le participant était invité à faire ressortir les points centraux de son expérience à travers une question dans laquelle il lui était demandé de formuler trois conseils à l'attention d'une ou un collègue qui envisagerait d'implanter une ÉTCSA dans le cadre d'un cours de première année.

Le schéma d'entrevue complet est présenté en annexe C. Un verbatim complet de l'entrevue a été réalisé. Toutefois, il est impossible de le rendre disponible en entier, car il contient un grand nombre d'informations qui permettraient d'identifier le participant : ainsi, pour des raisons de confidentialité, seuls des extraits choisis et anonymisés sont présentés dans le chapitre 4, et identifiés par la lettre E, suivis du ou des numéros de ligne correspondantes dans le verbatim.

Au niveau du déroulement, l'entrevue a une durée de 1h22m10s et elle a été enregistrée sur support audio dans un format AAC, enregistrement qui sera conservé selon les modalités prévues dans le certificat d'approbation éthique émis par l'institution d'enseignement de Julien (voir le formulaire de consentement, en annexe A).

3.3.2. *Le questionnaire*

Immédiatement avant la réalisation de l'entrevue, un questionnaire a été remis à Julien, qui l'a complété par écrit, de manière autoadministrée, en présence de la chercheuse. Ce questionnaire est composé de 17 questions fermées ou à court développement, regroupées en trois parties : a) le parcours professionnel, b) les expériences d'implantation d'ÉTCSA, et c) le profil sociodémographique. Le questionnaire complet peut être consulté à l'annexe B.

Les données recueillies à l'aide de ce questionnaire sont utilisées en complément des entrevues, qui demeurent la source principale de données, ce qui permet une triangulation des méthodes et facilite l'objectivation des perceptions rapportées par le participant, notamment grâce aux questions recourant à des échelles de réponses. Par souci de concision, lorsqu'il est fait référence à une question du questionnaire, celle-ci est désignée par la lettre Q, suivi du numéro de la question.

3.3.3. *Le dossier d'artefacts*

Le dossier d'artefacts est composé des documents utilisés par Julien dans le cadre de l'élaboration, de l'administration et de l'évaluation de l'ÉTCSA et a été fourni par celui-ci le jour de la réalisation de l'entrevue. Il contient 18 documents, qui ont été classés selon par type de matériel et par type de document. Au niveau du type de matériel, les catégories sont : a) matériel destiné aux étudiantes et étudiants, b) matériel utilisé uniquement par l'enseignante ou l'enseignant, et c) matériel utilisé par les étudiantes et les étudiants et l'enseignante ou l'enseignant. Au niveau du type de

document, les catégories sont : a) énoncé de travail ou consignes; b) Documents soutenant l'apprentissage (textes, notes de cours, documents présentant des notions, etc.), c) matériel d'évaluation, et d) autre. Le contenu du dossier a été compilé et classifié à l'aide de la Grille de collecte des artefacts, présenté en annexe D. Par souci de concision, lorsqu'il est fait référence à un artefact, celui-ci est désigné par la lettre A, suivi du numéro de document auquel il correspond dans la Grille de collecte des artefacts, présentés à l'annexe D.

Une des limites du dossier d'artefact est qu'il a été constitué a posteriori : en effet, lors de la conception et de la mise en œuvre de son ÉTCSA, Julien ne savait pas qu'il témoignerait un jour de son expérience dans une recherche. La principale répercussion au niveau de la composition est que les dix-huit documents remis sont essentiellement des documents de consignes ou d'évaluation des productions étudiantes. Ceux-ci permettent donc de documenter la version actuelle de son ÉTCSA, et uniquement dans une portion des phases d'administration et d'évaluation; aucun document ne témoigne de la phase d'élaboration, ni des différents ajustements réalisés au fil des expériences de mise en œuvre.

Par conséquent, le dossier d'artefacts a été utilisé pour confirmer les propos de Julien, d'illustrer concrètement certains exemples donnés par celui-ci et pour repérer les différentes références faites par celui-ci lors de l'entrevue réalisée. Ainsi, dans la présentation et l'analyse des résultats, les références au dossier d'artefact auront pour fonction l'illustration ou la confirmation de propos de Julien.

3.4. La méthode d'analyse et de traitement des données

La présente recherche est de nature qualitative/interprétative. Selon Paillé et Mucchielli (2012) :

L'analyse qualitative [...] peut être définie comme une démarche discursive de reformulation, d'explication ou de théorisation de

témoignages, d'expérience ou de phénomènes. La logique à l'œuvre participe de la découverte et de la construction de sens. Elle ne nécessite ni comptage ni quantification pour être valide, généreuse et complète, même si elle n'exclut pas de telles pratiques. Son résultat n'est, dans son essence, ni une proportion ni une quantité, c'est une qualité, une dimension, une extension, une conceptualisation de l'objet (p.11).

C'est précisément dans cette approche de la recherche que nous inscrivons notre démarche, car notre objectif n'est pas de quantifier la récurrence des facteurs favorisant ou défavorisant l'implantation d'ÉTCSA, mais plutôt d'en saisir la nature et les mécanismes dans un contexte donné.

De plus, étant donné les considérations méthodologico-épistémologiques présentées jusqu'à maintenant, l'approche qualitative s'impose. Tout d'abord, le caractère exploratoire de cette recherche, la volonté de documenter en profondeur les pratiques évaluatives étudiées, la nature hétérogène de ces pratiques et les difficultés de recrutement à grande échelle ne permettent pas de prélèvement quantitatif fiable. Ensuite, le besoin de comprendre la dynamique du phénomène à travers la compréhension que ses acteurs en ont, plutôt que d'étudier sa récurrence, commande une approche qualitative. Or,

[L]a démarche souple et émergente de la recherche qualitative/interprétative permet au chercheur de comprendre, de l'intérieur, la nature et la complexité des interactions d'un environnement déterminé, et d'orienter sa collecte de données en tenant compte de la dynamique interactive du site de recherche (Karsenti et Savoie-Zajc, 2011, p. 126)

Finalement, «[L]e deuxième facteur qui légitime la pertinence de la recherche qualitative/interprétative en éducation repose sur une qualité essentielle et intrinsèque à sa nature, à savoir tenir compte des interactions que les individus établissent entre eux et avec leur environnement.» (Karsenti et Savoie-Zajc, 2011, p. 125) Ceci est en droite ligne avec notre positionnement épistémologique.

De plus, étant donné le caractère exploratoire de notre recherche et le recours à l'étude de cas, notre démarche de traitement et d'analyse de données a suivi une logique inductive modérée. En recherche, l'induction est un «processus d'exploration qui s'attarde à l'observation de faits particuliers en vue d'y percevoir éventuellement une logique d'ensemble.» (Legendre, 2005, p. 770). Concrètement, les données spécifiques recueillies à travers l'étude de cas étudiée ont été décortiquées afin de dégager des éléments signifiants susceptibles de nourrir la formulation d'hypothèses quant aux conditions générales favorisant et défavorisant l'implantation d'ÉTCSA dans les cours de première année du DEC en sciences humaines.

Il existe plusieurs types de démarches inductives, parmi lesquelles nous avons retenu l'approche inductive délibératoire, qui «consiste à utiliser le cadre théorique comme un outil qui guide le processus de l'analyse [...] la grille d'analyse initiale [pouvant] toutefois être enrichie si d'autres dimensions ressortent des données.» (Karsenti et Savoie-Zajc, 2011, p. 138)

Ainsi, nous avons suivi le cycle de l'analyse de données qualitatives en 3 étapes -a) collecte des données; b) réduction des données; c) interprétation et vérification, lesquelles sont entrecoupées par un retour constant aux données (visualisation des données) - proposé par Karsenti et Savoie-Zajc à partir de l'ouvrage de Miles et Huberman (dans Karsenti et Savoie-Zajc, 2011, p. 139).

Finalement, pour les phases de réduction et d'interprétation des données, nous avons utilisé la catégorisation et interprétation des données selon des thèmes communs, une méthode d'analyse des données adaptées à l'étude de cas selon Fortin (2010). Le codage des données a été réalisé à l'aide du logiciel *QDA miner lite*, et le schéma de codage comprenait une double typologie : la première reflétant la structure de l'entrevue semi-dirigée, la principale méthode de collecte, et la seconde basée sur les catégories de problèmes recensés dans la littérature à propos de l'ÉSA, lesquelles ont été présentées dans le chapitre 2 et dans le tableau 1 - Synthèse des problèmes reliés à

la conception et la mise en œuvre d'une ÉSA recensés dans la littérature (Annexe E). Ce double codage des données s'inscrit parfaitement dans le respect du principe de visualisation des données, présenté ci-dessus.

4. LES ASPECTS ÉTHIQUES

Faire de la recherche avec des êtres humains exige un haut niveau éthique tout au long de la démarche, de l'élaboration à la diffusion. Les principes éthiques en vigueur dans la communauté scientifique ainsi que la Politique de recherche avec des êtres humains (PREH) en vigueur dans notre institution d'affiliation, le Cégep Édouard-Montpetit (Collège Édouard-Montpetit, 2013), ont guidé l'ensemble du processus. Voici plus précisément comment la présente démarche de recherche s'est inscrite dans ces grands principes :

4.1. Respect de la dignité humaine

«Ce principe, qui vise à protéger les intérêts multiples et interdépendants de la personne, depuis son intégrité corporelle jusqu'à son intégrité psychologique ou culturelle, constitue le fondement des obligations éthiques précisées ci-dessous.» (Collège Édouard-Montpetit, 2013, p. 9) En plus des enjeux liés à la confidentialité des propos, les outils de collecte des données ont été conçus afin que l'interaction entre la chercheuse et le participant soit empreinte de respect tout au long de la démarche.

4.2. Respect du consentement libre et éclairé

Chaque personne est réputée avoir le droit et la capacité de prendre des décisions libres et éclairées. Un consentement est *libre* lorsqu'il est donné sans coercition, influence indue ni manipulation. Il est *éclairé* quand le sujet dispose de suffisamment d'information sur la recherche et qu'il comprend celle-ci. (Collège Édouard-Montpetit, 2013, p. 9)

Lors du recrutement, la recherche a été présentée avec transparence. Toute l'information pertinente a été divulguée à une potentielle participante ou un potentiel

participant, pour qu'ils connaissent les objectifs poursuivis par la recherche et le type de traitement et de diffusion des résultats prévus. Un formulaire de consentement reprenant ces informations a été remis au participant sélectionné (voir Annexe A), lequel l'a signé avant le début de sa participation. Celui-ci conserve un droit de retrait tout au long de la démarche, sans pression de la part de la chercheuse advenant qu'il se prévale de cette option.

4.3. Respect des personnes vulnérables

Étant donné que la présente recherche ne concerne pas cette catégorie de personne, il n'y a pas d'enjeu relié au respect de ce principe de la PREH.

4.4. Respect de la vie privée et des renseignements personnels

Afin de respecter ce principe, les mesures suivantes ont été appliquées dans le cadre de la démarche :

1. Confidentialité de certaines informations sociodémographiques : Le cégep de provenance du participant, ainsi que la discipline enseignée, ne sont pas identifiés dans les documents qui seront diffusés publiquement, afin d'éviter l'identification par croisement d'information. Un pseudonyme (Julien) a été attribué pour la présentation des résultats et tous les documents d'analyse et de diffusion utilisent ce prénom fictif afin d'éviter la divulgation accidentelle de l'identité de l'interviewé. Le seul document sur lequel l'information nominative apparaît est le formulaire de consentement;

2. Le respect de la justice : Ce principe se définit comme suit :

La notion de justice réfère aux concepts d'impartialité et d'équité. Dans le domaine de la recherche, cette notion s'applique, d'une part, au processus d'évaluation des protocoles de recherche, qui doit faire appel à des méthodes et règles équitables et doit se dérouler de façon indépendante. Le

principe de justice s'applique aussi à la répartition des bienfaits et fardeaux liés à la recherche, de façon à ce qu'aucun segment de la population ne subisse plus que sa juste part de ses inconvénients, ni ne soit systématiquement privé de ses bienfaits. (Collège Édouard-Montpetit, 2013, p. 9)

Au niveau de la rigueur du processus de recherche, ceci sera assuré par la supervision institutionnelle de l'ensemble de la démarche de recherche – au premier plan par la directrice d'essai. Pour ce qui est de l'aspect de l'équilibre des bienfaits et des fardeaux liés à la recherche, comme la recherche est non expérimentale, il n'y a pas d'enjeux de fond liés à cet élément de ce principe;

3. *L'équilibre des avantages et des inconvénients et la réduction des inconvénients :*

«L'éthique moderne de la recherche exige que les inconvénients prévisibles d'un projet de recherche ne soient pas plus importants que les avantages escomptés, en vertu du principe de préoccupation pour le bien-être des sujets de recherche» (Collège Édouard-Montpetit, 2013, p. 9). Dans le cadre de la présente recherche, dans la mesure où les principes de respect de la vie privée, de respect de la confidentialité et de rigueur dans la démarche scientifique sont respectés, la possibilité que l'équilibre avantages/inconvénients de la recherche soit rompu est plutôt mince. Le seul danger possible est une récupération potentielle des résultats de recherche à des fins politiques par des intervenants extérieurs, mais ce risque est faible et indépendant des démarches de la chercheure;

4. *L'optimisation des avantages :*

L'optimisation des avantages réfère au devoir de viser le bien d'autrui et, d'un point de vue éthique, d'optimiser les avantages nets des projets de recherche, avantages qui peuvent profiter aux sujets eux-mêmes, à d'autres personnes ou encore, dans la majorité des cas, à la société et à l'enrichissement des connaissances. (Collège Édouard-Montpetit, 2013, p. 10)

La diffusion des résultats de la recherche assure le respect de ce principe;

5. *La minimisation des risques* : Le principe de minimisation des risques sera également respecté, notamment grâce à l'anonymisation des propos :

La recherche se situe sous le seuil de risque minimal lorsqu'on a toutes les raisons de penser que les sujets pressentis estiment que la probabilité et l'importance des éventuels inconvénients associés à une recherche sont comparables à ceux auxquels ils s'exposent dans les aspects de leur vie quotidienne reliés à la recherche » (Collège Édouard-Montpetit, 2013, p. 8)

Finalement, le respect de ces principes est garanti par la supervision professionnelle de la directrice de cet essai. Un Certificat d'approbation éthique a également été émis en date du 16 mai 2016, autorisant la collecte de données présentée précédemment.

5. LES MOYENS PRIS POUR ASSURER LA SCIENTIFICITÉ

La scientificité d'une recherche qualitative peut être évaluée à partir de quatre critères de rigueurs : la crédibilité, la transférabilité, la fiabilité et la confirmabilité (Fortin, 2010; Karsenti et Savoie-Zajc, 2011) Dans le cadre de cet essai, voici les mesures prises afin d'atteindre un niveau adéquat pour chacun de ces critères.

5.1. La crédibilité (aussi appelée validité interne):

Selon Gauthier (2010),

[l]a validité interne est la caractéristique d'une structure de preuve qui fait que les conclusions sur la relation de cause à effet reliant le facteur déclenchant au changement d'état de la cible sont solides et qui assure que les changements ne sont pas causés par la modification d'autres variables» (p. 191)

Il s'agit donc de «vérifier la plausibilité de l'interprétation du phénomène étudié.» (Karsenti et Savoie-Zajc, 2011, p. 140)

Selon Gauthier (2010) et Karsenti et Savoie-Zajc (2011), dans le cadre d'une étude de cas, deux stratégies clés contribuent à réduire l'influence des biais d'interprétation et favorisent l'approfondissement de l'analyse, donc augmente la crédibilité des conclusions de la recherche: 1) la multiplication des méthodes et des sources (aussi appelée triangulation) et 2) la tenue d'un journal de bord. Ce sont les outils qui ont été utilisés pour augmenter la rigueur dans le cadre de la présente recherche. Voyons quels sont leurs bénéfices respectifs et la façon dont ils ont été intégrés dans notre démarche :

1. *La triangulation des méthodes* : La multiplication des angles d'observation de l'objet à l'étude permet au chercheur de réduire les erreurs de mesures et de «comblent les lacunes ou biais de chacune des méthodes ou des sources d'information dont il fera usage.» (Gauthier, 2010, p.218) et d'«objectiver la démarche de coconstruction des savoirs» (Karsenti et Savoie-Zajc, 2011, p.144), ce qui s'inscrit dans la droite ligne du paradigme interprétatif. La triangulation des méthodes a été réalisée en recueillant les données auprès des participants à l'aide de trois méthodes de collecte : l'entrevue (méthode principale), l'analyse de corpus (le dossier d'artefacts) et le questionnaire.

2. *Le journal de bord* :

Pour Gauthier (2010), le journal de bord est un outil fondamental de la recherche :

Le caractère libre et privé du journal de bord lui permettra de tout noter, sans gêne ou discrimination. [...] il permet au chercheur de prendre conscience de ses biais et, surtout, de prendre des mesures spéciales pour vérifier ses observations lorsque celles-ci concordent avec ses idées de départ. Peut-être plus important encore, le journal de bord aide le chercheur à tisser des liens entre différentes observations ou différentes sources de mesures. Le chercheur saura tout noter et, surtout, se relire plusieurs fois pour atteindre une meilleure compréhension de la situation. [...] Bref, en notant tout dans son journal, le chercheur finira par prendre

conscience de ses biais et par objectiver sa pensée et ses interprétations. Ainsi, le journal l'aidera à accroître à la fois la validité de ses observations et la profondeur de ses interprétations. (p. 218-219)

Un journal de bord a été tenu tout au long de la démarche. Son contenu est très variable, allant des notes générales et réflexions personnelles aux questions et ébauches d'explication émises ponctuellement, sans oublier la documentation des difficultés rencontrées dans les différentes étapes de la recherche.

5.2. La transférabilité (aussi appelé validité externe)

Pour Gauthier (2010) : « la validité externe est la caractéristique d'une structure de preuve qui fait que les résultats obtenus sont généralisables au-delà des cas observés pour les fins de l'étude. » (p. 191) Il identifie 8 menaces à la validité externe d'une recherche : l'autosélection, l'effet de contagion, le contexte, les conditions expérimentales, les relations causales ambiguës, la réactivité aux prétests, le désir de plaire, le biais de l'analyste (les premiers ayant prise sur l'état de la source d'information et la seconde influençant la mesure des informations recueillies). Dans le cadre de la présente recherche, les biais à la validité externe les plus à craindre sont les suivants :

1. *L'autosélection découlant d'un échantillonnage non aléatoire* : « lorsque la sélection des individus à l'intérieur des groupes stratégiques est non aléatoire, les caractéristiques des individus sélectionnés peuvent être la cause de l'état ultérieur de la cible plutôt que le facteur déclenchant. » (Gauthier, 2010, p. 194) Du fait que l'échantillon ne comporte qu'un participant, ce risque est réel.
2. *Le désir de plaire* : l'utilisation de l'entrevue comme méthode de collecte principale et de l'identité de la chercheuse (elle aussi professeure en sc. humaines) qui y sont propices.

3. *Le biais de l'analyste* : Gauthier (2010) met en garde contre le danger de laisser les préconceptions influencer la démarche d'analyse du chercheur, malgré sa bonne foi. «Il est possible que les résultats obtenus soient davantage représentatifs des attentes du chercheur que de la réalité objective, sans qu'il ne cherche consciemment à biaiser les conclusions.» (Gauthier, 2010, p. 194)

La prise de conscience de ces pièges est le premier pas vers leur évitement. De plus, la présente démarche de recherche se déroulant dans le cadre de la production d'un essai de maîtrise, il y a eu vérification externe (ou débriefing) de la démarche et des analyses produites tout au long de la recherche à travers les rencontres avec la direction d'essai ainsi qu'avec des pairs dans le cadre des différents séminaires de maîtrise. Fortin (2010) définit cette technique ainsi : «Le *débriefing* par les pairs implique que le chercheur sollicite des échanges avec des personnes expérimentées dans les méthodes ou le contenu et qu'il sollicite les opinions des participants sur la crédibilité des interprétations élaborées par le chercheur et sur les résultats obtenus.» (p. 285) Cette approche a permis de réduire les risques associés aux menaces à la validité externe de la démarche et des résultats.

5.3. La fiabilité

Une étude est fiable lorsqu'il y a stabilité des données et constance des résultats :

On parle de fiabilité quand la répétition de l'étude avec les mêmes sujets dans des circonstances similaires produirait des résultats consistants. Dans la recherche de la fiabilité des données, le chercheur doit assurer l'exactitude des données et être en mesure de clarifier les schémas de codes et le processus d'analyse des données. Il fera appel à des personnes indépendantes pour vérifier si les résultats, les interprétations et les conclusions sont solidement appuyés. (Fortin, 2010, p. 285)

Les mesures prises pour augmenter la fiabilité de nos résultats découlent principalement des différentes triangulations effectuées (présentées précédemment).

Le processus de retour auprès du participant, advenant un besoin de précision, permet aussi de favoriser la fiabilité de l'étude en réduisant les possibles mésinterprétations des propos de la personne. De plus, la triangulation du chercheur par le biais d'une vérification externe par la direction d'essai et par les évaluateurs renforce cette quête de fiabilité. Finalement, la tenue d'un journal de bord a aidé la chercheuse à assurer une cohérence dans l'évolution de la démarche.

5.4. La confirmabilité

Fortin (2010) décrit ainsi le critère de la confirmabilité :

Le critère de confirmabilité (neutralité) réfère à l'objectivité dans les données et leur interprétation. Ce critère vise à s'assurer que les résultats reflètent bien les données et non le point de vue du chercheur. Les significations qui émergent des données doivent être vérifiées pour en évaluer la vraisemblance, la solidité et la certitude, de sorte que deux chercheurs indépendants obtiendraient des significations similaires à partir des données. Ces vérifications se font à rebours en faisant appel à d'autres personnes pour juger de la qualité des résultats obtenus. (p. 285)

Dans le cas de cette recherche, les différentes triangulations des méthodes et des analyses présentées précédemment favorisent la confirmabilité. De plus, cette démarche se déroulant dans le cadre de l'obtention d'un diplôme de maîtrise, ce critère de validité est aussi assuré par le comité d'évaluation de l'essai. Les résultats seront aussi soumis au regard et aux commentaires de la communauté scientifique lors d'éventuelles présentations dans des colloques et publications d'articles présentant la démarche et les résultats de cette recherche.

QUATRIÈME CHAPITRE

LA PRÉSENTATION ET L'INTERPRÉTATION DES RÉSULTATS

Le présent chapitre présente les résultats de la collecte de données ainsi que leur analyse et interprétation. La première partie sera consacrée à la présentation des données collectées à partir de l'entrevue, du questionnaire et du dossier d'artefact. Tout comme pour la structure de l'entrevue, la structure de présentation des résultats repose sur la typologie de Hardy (1996) selon laquelle la mise en œuvre d'une ÉSA se déroule en trois phases : la conception, l'administration et l'évaluation. En plus de l'objectif de recherche principal, cette première section poursuit également deux des objectifs spécifiques de recherche, soient a) identifier les enjeux et difficultés rencontrées par une enseignante ou un enseignant ayant expérimenté la mise en œuvre d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans le cadre d'un cours de première année du programme d'études préuniversitaires en sciences humaines, et b) identifier des conditions et actions perçues comme favorables à la mise en œuvre de pratiques évaluatives s'appuyant sur les principes d'une évaluation en situation authentique par un enseignant ou une enseignante ayant expérimenté la mise en œuvre d'une telle évaluation dans le cadre d'un cours de première année du programme d'études préuniversitaires en de sciences humaines.

Dans la seconde partie de ce chapitre, ces données sont analysées et interprétées à partir des éléments recensés dans la littérature portant sur l'ÉSA, afin d'atteindre le dernier objectif spécifique de recherche, soit analyser les enjeux et difficultés rencontrées et les conditions et actions perçues comme favorables identifiés par une enseignante ou un enseignant durant l'implantation d'une évaluation terminale

de cours en situation authentique dans le cadre d'un cours de première année du programme d'étude préuniversitaire en sciences humaines.

1. PRÉSENTATION ET ANALYSE DES RÉSULTATS

Il ressort de l'ensemble des données recueillies auprès de Julien qu'il considère que la mise en place d'une nouvelle ÉTCSA est très exigeante en temps et en ressource, mais que cela en vaut l'effort. Globalement, Julien est très satisfait de sa nouvelle approche évaluative, et son bilan est largement positif. En effet, dans le questionnaire qu'il a complété, il s'est déclaré «fortement en accord» lorsqu'interrogé sur son degré d'accord avec l'affirmation suivante : «La mise en place d'une ÉTCSA a eu un effet positif sur ma relation pédagogique avec les étudiants et étudiantes» (Q13a). Il s'est également déclaré «assez en accord» avec les affirmations suivantes :

- 1) «La mise en place d'une ÉTCSA a eu un effet positif sur les résultats (notes) des étudiants et étudiantes» (Q13b)
- 2) «La mise en place d'une ÉTCSA a eu un effet positif sur la qualité des productions (travaux) des étudiants et étudiantes» (Q13c)
- 3) «La mise en place d'une ÉTCSA a eu un effet positif sur la qualité des apprentissages réalisés par les étudiants et étudiantes» (Q13d)

Pour lui, un gain central de son approche actuelle est que « [ç]a me permet d'être beaucoup [...] en rétroaction», et cela a un impact direct sur la qualité des apprentissages réalisés.

Sur le plan de l'impact sur le développement professionnel, il fait également un bilan positif : Julien se déclare «assez en accord» avec l'affirmation selon laquelle la mise en place d'une ÉTCSA a eu un effet positif sur son développement professionnel (Q13e), et il a l'intention de poursuivre dans cette voie.

Toutefois, il se dégage aussi certains constats quant aux difficultés associées à une approche de type ÉTCSA. Il considère aussi que mettre en place une ÉTCSA est plus complexe et difficile qu'un enseignement traditionnel, et est plus exigeant pour l'enseignante ou l'enseignant. Pour ces raisons, il le déconseille à une personne novice.

De façon très simple: un prof, avec le matériel que j'ai monté, un prof nouveau qui arriverait ici et qui me disait "je veux faire ça cette session-ci" je dirais "non, ce n'est pas une bonne idée, commence de façon plus traditionnelle, commence par maîtriser ton contenu" Après, oui, mais attend un peu [...] parce que la mise en application, la première fois, est plus déstabilisante. C'est beaucoup plus simple de faire, comme je faisais avant, et comme je fais dans d'autres de mes cours, de dire "tu vas faire un travail de 5 pages, de 10 pages sur X et remet-moi ça dans 3 semaines"... que [d'opter pour cette nouvelle forme d'évaluation]. Il y a beaucoup plus de consignes... je prends beaucoup plus de temps en classe pour expliquer des consignes d'évaluation que ce que je faisais avant. [...] Tout ça pour dire qu'au début, la côte est quand même plus abrupte à monter, qu'il faut être un peu plus en contrôle de sa matière, de la pédagogie, avant d'aller dans une approche comme [la mienne]. [...] C'est juste que le défi est un peu plus grand pour commencer. Mais après ça, une fois que c'est lancé, après une première session [ça va] (E.328-378)

Voyons maintenant comme ces enjeux se manifestent pour chacune des phases de la mise en œuvre d'une ÉTCSA.

1.1. Phase d'élaboration

La première partie de l'entrevue ainsi que plusieurs questions du questionnaire ont porté sur la phase d'élaboration de l'ÉTCSA mise en œuvre par Julien. Lors de l'entrevue, cette phase a été définie comme

la phase de conception du projet d'évaluation en situation authentique et de rédaction du matériel pédagogique et des outils d'évaluation associés. Il s'agit donc de l'étape qui précède la présentation en classe et de la remise du matériel pédagogique et du matériel d'évaluation aux étudiantes et étudiants (Annexe C, p. 180).

Lorsqu'interrogé sur son vécu concernant la phase d'élaboration de son ÉTCSA, Julien a abordé spontanément les thèmes suivants : le temps, l'interaction avec des collègues, la disponibilité de ressources spécialisées et la latitude sur le plan institutionnel. Un résumé des informations présentées ci-dessous peut également être consulté en annexe F (Tableau 2 – Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'administration de l'ÉTCSA).

1.1.1. Le besoin de temps

Explicitement identifié par Julien comme le facteur le plus important pour la phase d'élaboration de son ÉTCSA, le besoin de temps est un thème récurrent tout au long de l'entrevue, ainsi que dans le questionnaire.

Dans son cas, Julien a bénéficié d'une subvention externe sous forme de libération ainsi que d'une libération attribuée par le cégep qui l'embauche pour l'élaboration de son projet d'ÉTCSA et du matériel pédagogique et évaluatif associé. Cela lui a permis de consacrer une portion substantielle de son temps de travail à cette tâche durant la réalisation de cette phase. C'est le facteur qu'il identifie comme le plus déterminant dans le succès de la phase d'élaboration de son projet : « Avoir le temps pour le faire. Ça, ça a aidé. C'était comme "ça" le défi. » (E.65-66). Selon Julien, il n'est pas réaliste d'élaborer ce genre de projet dans le cadre d'une tâche normal d'enseignement. Lorsqu'interrogé sur la manière dont il a réussi à contourner cette difficulté, il répond : « je suis allée me chercher du dégrèvement. Pis, c'est vraiment pratico-pratique, c'est vraiment ça que j'ai fait, dans le sens où mon objectif initial, quand j'ai répondu à l'appel de projets du [nom de l'organisme], c'était de construire cette évaluation-là. (E.75-80)

Lorsqu'interrogé sur les trois conseils généraux à donner à quelqu'un qui envisage développer un projet d'ÉTCSA, Julien identifie le fait de s'assurer d'avoir du

temps pour concevoir le projet comme le premier élément à considérer. Il est à noter que ce besoin de temps se fait particulièrement sentir dans la phase d'élaboration, selon les propos de Julien.

1.1.2. L'interaction avec des collègues

Pour l'ensemble de son projet d'ÉTCSA, Julien était le principal architecte du projet. Il travaillait principalement seul, bien qu'il ait consulté ponctuellement des collègues et des ressources spécialisées.

Le fait de travailler principalement seul sur ce projet a été une difficulté identifiée par Julien.

Ce qui a été compliqué aussi, c'est de le faire seul. Ça, ça a été quelque chose, un défi, pour chercher l'équilibre, pour réfléchir à tout ça.... [...] telle consigne versus telle consigne, est-ce que c'est équivalent ? Dans ma tête peut-être que oui, mais peut-être que dans la tienne, tu vas dire "il me semble que non". Des fois, j'ai pu aller voir les collègues [qui enseignent la même discipline] et leur dire "j'ai besoin de votre avis, aidez-moi un peu", mais étant donné que c'était mon projet, que c'est moi qui était dégrevé..... quelques fois, j'aurais aimé ça avoir une autre personne dans le projet avec moi pour qu'on puisse confronter nos idées, etc. (E.131-141)

Pour Julien, à l'étape de l'élaboration, les interactions avec les collègues permettent de valider la viabilité des certaines dimensions du projet d'ÉTCSA et de corriger le tir au besoin. Il a trouvé utile de pouvoir consulter ponctuellement ses collègues enseignant la même discipline que lui lors de l'élaboration du matériel pédagogique relié à l'ÉTCSA, afin d'obtenir leur opinion sur la convivialité de certaines consignes et documents et la faisabilité logistique de certaines dimensions du projet.

[ça m'a aidé de] jaser avec les collègues. J'ai des consignes que j'ai montrées à des collègues, qui m'ont dit [parfois] "il me semble qu'il manque quelque chose ici, il me semble que cette tâche-là est plus facile

que celle-là" "tu as peut-être raison, peut-être que si je rajoute cette consigne, cette exigence-là".... "ah oui" (E.703-707)

Selon lui, cette possibilité de consulter des collègues lui a été particulièrement utile afin d'assurer l'équité des évaluations sur le plan de leur conception, surtout du fait que les tâches de son ÉTCSA impliquent parfois des rôles et tâches dont certaines portions diffèrent entre les étudiantes et les étudiants d'un même groupe. Il identifie d'ailleurs cette possibilité de consulter les collègues comme le plus important rempart en matière d'équité des évaluations, au niveau de ce qui peut être contrôlé lors de la conception de l'ÉTCSA.

En plus de cette aide en cours de projet, l'interaction avec les collègues peut aussi être source d'inspiration. Dans son cas, une présentation faite par des collègues d'un autre département a contribué à faire émerger l'idée et une amorce de structure du projet d'ÉTCSA qu'il a finalement développé.

Il y a quelques sessions de cela, ça été un des éléments qui a mis une puce à mon oreille de faire mon projet comme je l'ai fait. On avait eu une journée pédagogique utile, et oui ça existe !, où c'était des profs qui avaient fait des ateliers sur différentes thématiques, et j'avais assisté à une présentation de deux de mes collègues en [nom d'une discipline], qui faisaient du travail d'équipe et qui disaient que justement dans l'évaluation, ils faisaient des consignes différentes selon les étudiants, dans la même perspective que moi. [Ils disaient que] parfois, ça prend plusieurs rôles, et qu'il y a une façon de rendre ça équitable. (E.805-814)

Pour Julien, cette interaction avec les collègues est suffisamment importante et porteuse pour qu'il l'identifie parmi les trois conseils qui à formuler à l'intention d'une personne souhaitant élaborer une ÉTCSA: «Si possible, de pouvoir consulter un collègue, pour qu'il puisse te donner des conseils.» (E. 962-963)

1.1.3. Disponibilité de ressources spécialisées

L'élaboration d'ÉTCSA peut être complexe et impliquer des savoirs et habiletés qui ne sont pas toujours pleinement maîtrisés par la professeure ou le professeur qui la conçoit. Les ressources spécialisées peuvent alors être d'un grand secours. C'est du moins l'expérience de Julien. Dans son cas, le projet d'ÉTCSA impliquait un volet technologique, qui visait à rendre viable la réalisation de son projet dans le cadre académique associé à un cours collégial.⁷ Ainsi, la présence et la disponibilité d'un conseiller pédagogique spécialisé en technologie de l'information et des communications ont été d'une grande aide, et clairement un élément facilitant dans la démarche de Julien.

Ce qui m'a aidé dans tout ça, c'est mon conseiller technopédagogique [...] qui a embarqué dans le projet, qui a dit "oui, moi je vais t'aider". On s'est rencontré à quelques reprises, il m'a aidé avec [l'intranet du cégep]. J'ai dit: "je veux faire telle chose avec [l'intranet du cégep], est-ce que je peux ?" Il allait se renseigner, il appelait [nom de la compagnie en charge du logiciel intranet du cégep]. (E.181-186)

Cet exemple illustre une situation où de telles ressources permettent d'aller chercher une expertise que la professeure ou le professeur ne possède pas, en plus de lui permettre de déléguer certaines tâches, pour lesquelles ces ressources spécialisées sont plus performantes.

Dans le cas de Julien, cela a été une aide précieuse, au point où il l'identifie comme un élément important de la réussite de son projet. Il faut toutefois être prudent quant à l'ampleur et à la nature de l'aide spécialisée nécessaire à un projet d'ÉTCSA du fait que les besoins de Julien étaient très reliés au contexte technologique particulier de son projet. Également, dans son cas, ce besoin s'est fait sentir principalement dans

⁷ Afin d'assurer l'anonymat, il n'est pas possible d'entrer dans le détail concernant la dimension technologique du projet d'ÉTCSA conçu par le participant.

la phase d'élaboration du projet, alors que pour d'autres formats d'ÉTCSA, ce besoin pourrait se faire sentir dans d'autres phases.

1.1.4. Le contexte institutionnel et académique

Le contexte institutionnel est également une dimension importante à considérer lors de la phase d'élaboration. Dans le cas de Julien, ce contexte a parfois été facilitant, parfois contraignant.

Au niveau des contraintes, il a dû faire modifier le plan-cadre de son cours. Toutefois, il insiste sur le fait que la démarche n'a pas été compliquée, et que cela n'a été une entrave à son projet :

J'ai fait modifier le plan-cadre du cours [...] La façon dont [le plan-cadre original] était formulé, mon projet pouvait plus ou moins rentrer. Mais quand j'ai présenté ça à mes collègues, ils ont dit "[il n'y a pas de problème]". Ça respectait l'esprit du plan-cadre, mais ça ne respectait pas la lettre. On a corrigé le plan-cadre, je pense que c'était quelque chose comme changer un "ou" pour un "et/ou", c'était quelque chose du genre.... [...] Mais ça a été comme une lettre à la poste. Il n'y a pas eu de débat. Ça a été de dire "voici la modification qu'on veut faire, voici pourquoi" et ok, c'est beau. (E.787-797)

Une difficulté rencontrée par Julien lors de la conception de son ÉTCSA est la prise en compte de certains éléments propres au contexte académique en général, et au contexte collégial en particulier, tout en maintenant l'authenticité de la tâche et de son contexte de réalisation, lesquels demandaient que les rôles attribués aux étudiantes et étudiants ne soient pas identiques, mais plutôt complémentaires. Or, cette variabilité des tâches ne pose pas problème dans le monde extrascolaire, du fait qu'elle résulte de la fonction de chacun. Toutefois, dans un contexte académique, cela pose des enjeux d'équité et d'équivalence, du fait que l'ÉTCSA a des répercussions sur la réussite ou non d'un cours, et ultimement, sur les dossiers d'admission à l'université des étudiantes et des étudiants.

La grande difficulté, ça a été d'équilibrer le tout. Surtout que dans ce cadre d'évaluation-là, tous les étudiants n'ont pas le même rôle, n'ont pas les mêmes tâches. De s'assurer qu'en termes d'évaluation, quand on ramène ça dans un contexte d'apprentissage collégial, de cote R, etc., que ce soit relativement équivalent, ça, ça n'a pas été facile. Ça a été ça le plus gros défi. Toi, si je te demande de faire un travail de 3 pages là-dessus, [et que toi] je peux pas te demander de faire ce travail-là à cause de ton rôle, mais que je dois te demander de faire 2 travaux de 1 page et demi, là je simplifie beaucoup [mon explication] mais.... ça c'était la plus grande difficulté. (E. 35-44)

Sur ce point, Julien mentionne également l'importante latitude pédagogique dont disposent les professeures et professeurs de son institution d'enseignement, latitude qui a, selon lui, permis l'innovation nécessaire à des projets d'ÉTCSA comme le sien. Dans son cas, le fait que la non-standardisation des rôles et tâches impliqués dans son projet d'ÉTCSA ne cause pas problème a permis l'existence même de ce projet et a facilité son élaboration. Il n'a pas eu à se justifier auprès de membres de l'administration ou de conseillers pédagogiques à ce propos, et associe cette latitude au fait que dans son institution, il y a une culture d'autonomie du travail professoral et une ouverture à la diversité des pratiques pédagogiques (E. 798-820).

Sur le plan des contraintes institutionnelles et académiques, l'autre élément abordé par Julien est celui relié au nombre d'heures d'étude et de temps en classe associés à un cours. Ces contraintes étaient centrales durant le processus d'élaboration afin que l'ÉTCSA soit réellement applicable dans le contexte académique. La solution trouvée afin de contourner cette difficulté est de diminuer l'authenticité de certaines dimensions de la tâche ou du contexte. Dans son cas, il a opté pour qu'une partie de la tâche se réalise en ligne, alors que dans le contexte extrascolaire, celle-ci se réaliserait en présence (diminution d'authenticité du contexte de réalisation) et a éliminé certaines dimensions de la réalité afin de simplifier la tâche dans son ensemble. « C'est plate, mais j'ai le temps de faire tout ça en classe, j'ai pas le temps (...) c'est tout simplement ça ! » (E.280-282)

Finalement, Julien mentionne l'importance de maintenir une latitude individuelle pour la professeure ou le professeur mettant en place une ÉTCSA, au niveau de la forme de celle-ci, afin de pouvoir adapter facilement et rapidement le projet. Pour Julien, il est important que la professeure ou le professeur qui met en place une ÉTCSA puisse l'adapter à son contexte d'enseignement et à ses préférences : «[mon projet d'ÉTCSA,] je l'ai construit à ma sauce à moi, c'est ma personnalité qui est dedans. Donc, un autre prof, il devra réajuster certains trucs avec sa réalité à lui ou elle.» (E.374-376)

Un résumé des informations présentées ci-dessous peut également être consulté en annexe F (Tableau 2 – Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'administration de l'ÉTCSA).

1.2. Phase d'administration

La seconde partie de l'entrevue, et certaines sections du questionnaire portaient sur la phase d'administration de l'ÉTCSA. Cette phase a été définie comme

la phase durant laquelle la tâche est présentée aux étudiants, puis réalisée par ceux-ci. Cela inclut toutes les phases de réalisation, y compris les moments de rétroaction formative, et se termine à la remise de la production finale ou au moment de réalisation de la performance, selon le type de tâche demandée. (Annexe C, p.180)

Dans le questionnaire complété par Julien, plusieurs réponses mènent au constat que l'implantation d'une ÉTCSA a impliqué une augmentation de la tâche du professeur durant la phase d'administration. Il déclare notamment être «assez en accord» avec l'affirmation selon laquelle «la mise en place d'une ÉTCSA a occasionné une augmentation du nombre d'heures consacrées au soutien des étudiants et étudiantes

à l'extérieur des heures de cours» (Q13h), et l'affirmation selon laquelle «la mise en place d'une ÉTCSA a occasionné une augmentation de la tâche (explications, rétroactions, etc.) nécessaire à l'encadrement des étudiants et étudiantes à l'extérieur des heures de cours» (Q13j) Il déclare également être «assez en accord» avec le fait que «la mise en place d'une ÉTCSA a complexifié son travail de professeur durant les heures de cours» (Q13l), ainsi qu'à l'extérieur des heures de classe (Q13k). Il déclare également être «très en accord» avec l'affirmation selon laquelle «la mise en place d'une ÉTCSA a occasionné une augmentation de la tâche (explications, rétroactions, etc.) nécessaire à l'encadrement des étudiants et étudiantes à l'intérieur des heures de cours» (Q13i). Ainsi, les réponses au questionnaire permettent de constater une intensification du travail du professeur lors de la phase d'administration d'une ÉTCSA, comparativement à une approche plus traditionnelle de l'enseignement. Les motifs expliquant cette perception sont abordés en détail lors de l'entrevue réalisée avec celui-ci.

Premier constat, il y a une distinction à faire entre la première mise en œuvre d'une ÉTCSA et ses mises en œuvre subséquentes : en effet, les principales difficultés rapportées par Julien sont rencontrées lors de la première implantation, et une partie de celles-ci s'estompent lors des mises en œuvre subséquentes, au gré des mesures et adaptations mises en place pour pallier les problèmes rencontrés initialement. Ainsi, les propos de Julien concernant la phase d'administration de l'ÉTCSA peuvent être regroupés autour de deux pôles (chacun composé de plusieurs éléments) : les enjeux associés à la première mise en œuvre et les enjeux récurrents ou contextuels. Un résumé des éléments composant ces deux pôles est également présenté en annexe F (Tableau 3 – Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'administration de l'ÉTCSA).

1.2.1. Les enjeux associés à la première mise en œuvre

Les propos de Julien lors de l'entrevue permettent de distinguer deux types de difficultés propres à la première expérience de mise en œuvre de son ETCSA : les difficultés et besoins non anticipés des étudiantes et étudiants et la maîtrise imparfaite de la démarche et du matériel de la part de l'enseignante ou l'enseignant.

1.2.1.1. Les difficultés et besoins non anticipés des étudiantes et étudiants : Lors de la première implantation de son ÉTCSA, Julien a rencontré certaines difficultés non anticipées liées aux connaissances technologiques et institutionnelles des étudiantes et des étudiants, qu'il avait surestimés :

Très banal: je leur fournissais un calendrier: de telle date à telle date vous devez faire tel nombre d'interventions.... mais ce ne sont pas tous les étudiants en première année qui comprennent les tableaux comme il le faut. [...] J'ai beaucoup de fichiers que je mets sur LÉA. Même ça, au début, les étudiants, ça les mélangeait un peu [parce qu']ils sont en apprentissage technologique en même temps [...] Juste en termes d'évaluations, [les étudiants et étudiantes se disaient] quelles évaluations s'adressent à moi ? Donc au début, ç'a été un peu mélangeant pour les étudiants. Mais bon. (E.202-223)

Pour Julien, la première mise en œuvre a nécessité d'être à l'écoute des étudiantes et des étudiants afin d'identifier leurs difficultés : «des fois, entre ce que je leur enseignais et le niveau d'exigence que j'avais..... oups ! Peut-être que ça va un petit peu trop vite pour eux. Notamment dans un contexte de première année. Parfois certains ne suivaient pas.» (E.195-198) Pour surmonter ces difficultés, il a réorganisé la séquence d'introduction de la tâche d'ÉTCSA : alors que dans la première mouture, il avait conçu le projet pour mettre les étudiantes et les étudiants dans le bain dès le départ, maintenir une interaction constante entre les nouvelles notions et le déroulement de l'activité, après quelques mises en œuvre, il a décidé de séparer davantage la théorie et l'application.

Je voyais l'explication [des contenus associés au cours] au début de la session pour donner les outils. Mais assez rapidement, je rentrais dans [la tâche associée à l'ÉTCSA], effectivement, et là ça posait un peu plus de problèmes. Là c'est un peu plus tard, donc j'ai gardé l'ordre de présentation de la matière théorique, mais c'est un peu plus tard qu'on commence [les tâches associées à l'ÉTCSA], donc ils ont un peu plus le temps d'assimiler, de juste comprendre comment fonctionne... le collégial à la limite, pour les étudiants de première session. De s'acclimater au prof, la dynamique du cours. (E. 285-295)

1.2.1.2. La maîtrise imparfaite de la démarche et du matériel de la part de l'enseignante ou l'enseignant : À ces besoins non anticipés chez les étudiantes et les étudiants s'ajoutent ceux associés à la maîtrise imparfaite de la démarche et du matériel de la part de la professeure ou du professeur, inhérente à la première expérience de mise en œuvre : «la première fois qu'on donne un cours, on est un peu, je dirais pas brouillon.... disons pas 100% en maîtrise. Là, c'était la première fois que j'utilisais ces évaluations dans une perspective quand même complexe.» (E.188-190) Bien que cette difficulté soit présente pour tous les types de cours et d'évaluation, Julien identifie cette difficulté comme plus importante dans le cadre de son expérience d'ÉTCSA, du fait de la moins grande standardisation d'une telle approche et des résultats anticipés. Ainsi, Julien a aussi fait face à des dimensions du matériel ou de la tâche qui n'avaient pas été anticipées lors de l'élaboration. Il donne l'exemple du traitement d'un questionnaire dans lequel il demandait aux étudiantes et aux étudiants d'identifier leurs champs d'intérêt afin de les assigner à des fonctions qui correspondant le mieux à ceux-ci : s'il avait prévu ledit questionnaire, il n'aurait pas anticipé la complexité du traitement de cette information à l'échelle d'une classe.

Je me suis dit que j'allais construire un questionnaire que je distribue aux étudiants pour voir leurs intérêts. [...] Mais pratico-pratique, comment je traite ces données-là. Il faut que je me monte un fichier Excel, comment je vais articuler ça, comment je mets ça en place. Donc vraiment la mise en application. J'ai fait "ok, il y a quand même du travail là !". Donc, ça a été quand même un défi. (E.170-179)

On retrouve dans le dossier d'artefacts des traces de ces outils élaborés par Julien pour résoudre les difficultés rencontrées, dont les documents A1, A2, A3, A5 et A17 (voir annexe D pour la liste des artefacts). L'étude de ces artefacts a permis de valider les propos rapportés en entrevue concernant l'ampleur et la précision des outils que le participant a dû construire pour faire face à ces différents enjeux logistiques.

Pour faire face à ces difficultés, Julien insiste sur l'importance d'adopter une grande flexibilité, d'être ouvert à adapter le déroulement en cours de route, et en vue des sessions subséquentes, et accepter qu'il y aura parfois, tout particulièrement lors des premières mises en œuvre, des impacts sur la validité, la fidélité, voire l'équité, des évaluations, le temps de faire les ajustements. :

[la première fois] je leur ai dit au début de la session "bon, vous allez être un groupe cobaye".... j'ai été moins sévère aussi, il y a des choses qui ont pu passer. [...] [il a fallu que j'arrive] à démêler le tout dans quelque chose qui est quand même passablement complexe, dans ce cadre-ci, ça a été un autre bon défi. (E.223-228)

Ainsi, à la suite de cette première expérience, Julien a réajusté une partie de son matériel, ajouter quelques explications afin de répondre aux besoins de tous les étudiantes et les étudiants, comme en fait foi l'étude certains artefacts fournis, notamment le document A17, présentant un calendrier détaillé des étapes afin de guider les étudiantes et les étudiants, et le document A3, présentant la répartition des rôles au sein d'un groupe (voir annexe D pour consulter la liste des artefacts).

À cette attitude de flexibilité s'ajoute le besoin de disposer de temps afin de pouvoir faire face aux éléments non anticipés et effectuer les ajustements nécessaires.

Finalement, Julien mentionne qu'il a eu la chance de bénéficier d'un soutien supplémentaire en classe lors de la première implantation de son ÉTCSA, car il avait un stagiaire à ce moment.

Et pour m'aider, encore une fois, c'est le hasard [...]j'avais un stagiaire cette fois-là, que je connaissais qui avait déjà fait un stage informel avec moi avant de faire ce stage formel là dans le cadre de son programme.... donc ça, ça m'a aidé, parce que des fois, pendant que j'expliquais ce qui se passait, bien lui aussi il pouvait faire la même chose à d'autres équipes. Ça été utile d'avoir cette fois-là un deuxième appui, et ça a été un des facteurs de motivation à ce que je fasse ce projet cette fois-là. [...] ça m'a aidé d'avoir eu deuxième "prof" sur place. (E.317-324)

La présence de ce stagiaire a été un élément facilitant lors de la première implantation de l'ÉTCSA conçu par Julien, ce qui lui a permis de faire face aux enjeux déclarés dans le questionnaire et l'entrevue concernant la charge accrue de soutien aux étudiantes et aux étudiants en classe (Q13-i).

À la suite de cette expérience, il croit qu'étant donné les inévitables imprévus associés à la première implantation et la nécessité de faire des ajustements rapides, parfois à l'intérieur de courts délais, l'apport d'un soutien supplémentaire est un élément qui facilite le déroulement.

1.2.2. Les enjeux récurrents ou contextuels

En plus des enjeux propres à la première mise en œuvre de son ÉTCSA, Julien aborde certains enjeux et certaines difficultés propices à survenir dans les mises en application subséquentes. Celles-ci peuvent être regroupées en deux catégories.

2.2.2.1. Besoin accru de soutien et d'accompagnement de la part des étudiantes et des étudiants : Tel que rapporté dans le questionnaire (Q13h-i-j) et dans l'entrevue, Julien déclare que la mise en place de l'ÉTCSA qu'il a conçue a engendré une augmentation du soutien à donner aux étudiantes et aux étudiants, et de l'accompagnement de ceux-ci dans la réalisation de la tâche. Il associe cela en partie au fait qu'au début, les étudiantes et étudiants se sentent un peu désorientés et on a besoin d'accompagnement.

Plusieurs consignes en soi peuvent être mélangeantes... plus le volet technologique, plus les consignes qui sont liées à la matière, ça fait qu'il y a quand même un petit défi pour les étudiants, de s'approprier tout ça. Il faut juste prendre le temps de bien expliquer le tout. (E.410-414)

Pour faire face à ce besoin, Julien a dû prévoir plus de temps en classe pour accorder ce soutien.

Il faut prévoir du temps [...]. J'essaie de faire travailler le plus possible [en classe] la première fois que je donne mes consignes, d'être en labo informatique, de leur donner du temps pour faire leurs premières interventions. Honnêtement, je ne réussis pas très bien [...] à y arriver, à avoir assez de temps, mais bon..... mais l'idéal c'est ça. Qu'ils aient du temps en labo informatique, que je puisse circuler. [...] "Ma question, est-ce qu'elle est bonne ?" "oui, ça marche" "Est-ce que mon intervention est bonne ?" "oui, elle est bonne" "est-ce que ça, ça marche ?" "non ça ne marche pas, parce qu'il te manque si, il te manque ça" Ça c'est quelque chose qui rassure les étudiants, et qui permet de mieux partir. Je réussis à moitié à le faire présentement. (E.915-932)

La gestion du temps en classe et la priorisation entre les activités reliées à la tâche de l'ÉTCSA et la présentation de notions plus théoriques reliées au cours est un enjeu de taille auquel fait face Julien. Il rapporte également qu'il y a plus de gestion des interactions avec les étudiantes et les étudiants, notamment à l'extérieur de la classe, depuis l'implantation de son modèle d'ÉTCSA (Q13h et Q13j). Les modalités de communication, telles que la messagerie interne, aident à la fluidité, mais cela demande du temps pour l'enseignante ou l'enseignant. Selon Julien, cela s'explique notamment par le fait que le déroulement n'est pas aussi prévisible que dans un travail plus traditionnel, avec des consignes fixes et immuables. Aussi, le fait que tous les étudiantes et les étudiants ne font pas exactement la même chose réduit les possibilités d'entraide, même si au final, la démarche encourage la collaboration. Pour Julien, cela a entraîné une augmentation des communications avec les étudiantes et étudiants à l'extérieur de la classe, qu'il associe directement à l'introduction de l'ÉTCSA.

C'est plus de temps, parce que les étudiants....il y a plus d'étudiants qui sont perdus, au niveau de ce qui se passe, surtout au début. Surtout que ce n'est pas tous les étudiants qui ont les mêmes [tâches], ça fait en sorte qu'il ne peut pas toujours demander à son ami ce qu'il doit faire, parce que ça se peut que son ami n'ait pas à le faire. Il y a aussi un autre élément. Comme j'ai plus d'évaluations tout simplement, c'est à chaque semaine, bien j'ai plus d'occasions qu'un étudiant me pose une question. Pour un étudiant, s'il a tel travail à faire, un document de 5 pages à me remettre dans trois semaines, il va peut-être m'écrire une fois ou deux, déjà c'est pas pire, mais voilà. Tant dis que si tu as cette évaluation-là, puis telle autre, que quelqu'un me pose une question.... un cas simple [il raconte que parfois, certains étudiants font des erreurs dans une tâche impliquant une interaction avec d'autres étudiants, rendant impossible la poursuite de la tâche par ce dernier] Alors l'étudiant [...] m'envoie un message pour me demander "ben là, qu'est-ce que je fais ?" Je lui réponds [en lui indiquant comment réagir à la situation] [...] Il y a plein de situations où l'étudiant [se dit] " qu'est-ce que je fais avec ça ?" Mais en général, ça se répond simplement. Ça n'a pas trop alourdi ma tâche, disons, on n'est pas dans ce niveau là non plus. Mais oui, j'ai plus d'interactions, surtout par [nom de la messagerie interne]. (E.548-582)

Julien mentionne que c'est particulièrement le cas dans les cours de première année du programme de sciences humaines, dans lesquels une partie importante des étudiantes et des étudiants inscrits en est à sa toute première session ou année au cégep.

Pour faire face à ce besoin accru de soutien, Julien a modifié sa planification de cours afin de prévoir du temps en classe et à l'extérieur de la classe pour le soutien individualisé auprès des étudiantes et des étudiants. Il s'assure également de

rassurer les étudiants souvent.... [...] surtout au début de la session, au moment des premières explications, ils ne comprennent pas trop, parce qu'il y a plein de choses en même temps, il y a beaucoup de consignes à donner en même temps. Il faut aussi prévoir en classe ce temps-là justement. Surtout les premières fois, ça prend du temps. (E.910-915)

Il a également adapté le matériel pédagogique, notamment en élaborant des consignes très détaillées, dont on retrouve la trace dans le dossier d'artefacts, notamment les documents A6 à A16 (voir annexe D pour la présentation de la liste des

artefacts). Durant l'entrevue, il rapporte également avoir mis à la disposition des étudiantes et des étudiants des exemples, créés par lui ou inspirés des sessions antérieures, pour illustrer concrètement la tâche.

2.2.2.2. Variabilité du contexte pédagogique et les imprévus : Une des principales difficultés rapportées par Julien relève de la variabilité du contexte d'administration. Il donne comme exemple une variation importante au niveau du nombre d'étudiantes et d'étudiants inscrits dans un groupe : «un exemple très simple: quand j'ai rédigé [mon concept d'ÉTCSA], j'ai fait un exemple pour une classe de 30 étudiants. Si j'en ai [disons] 22. Ok, quels [rôles] je garde, quels je coupe, quels je fusionne ensemble ?» (E.68-170) Au-delà de la variabilité inhérente au contexte pédagogique, Julien a souvent dû faire face à des imprévus en cours de sessions. Il donne l'exemple d'un étudiant qui avait une tâche déterminante pour la suite du projet pour l'ensemble de la classe et qui a soudainement abandonné le cours, sans avertir : or, ceci a eu des répercussions sur le travail des autres étudiantes et étudiants, et Julien a dû faire face à cette situation au début d'un cours lors duquel plusieurs étudiantes et étudiants devaient utiliser la présentation faite par l'étudiant ayant quitté pour poursuivre leurs tâches respectives, se retrouvant alors bloqués dans la poursuite de leurs tâches.

Devant de telles situations, Julien a eu des réactions parfois ponctuelles, parfois structurelles. Les solutions ponctuelles étaient principalement des réorganisations des consignes sur-le-champ, afin de permettre la poursuite du projet, malgré la perte d'authenticité du contexte que cela engendrait parfois. Sur le plan des solutions structurelles, elles visaient principalement à s'adapter au contexte institutionnel : par exemple, devant les problèmes chroniques d'abandon de certaines étudiantes et certains étudiants dans le premier mois de la session, Julien a choisi de repousser le début de la tâche à un moment où les groupes étaient plus stabilisés. « Donc de commencer plus tard, il y a moins de réajustements à faire». (E.302-303)

Pour faire face à ces enjeux, Julien insiste sur l'importance de disposer de la latitude nécessaire lors de prises de décisions. Pour Julien, si la collaboration entre collègues est précieuse lors de la phase d'élaboration, la latitude nécessaire à la prise de décision rapide et spontanée est essentielle pour faire face aux défis de la phase d'administration. Pour Julien, une des clés du succès est de savoir s'adapter : «Et d'avoir une bonne flexibilité aussi, dans le sens d'être capable de réagir en fonction de ce qui se passe durant la session, et corriger le tir, se réajuster, être prêts [...] [Quand] il se passe quelque chose, il faut être capable de corriger le tir. » (E.768-771)

Cette latitude lui est nécessaire pour faire face aux imprévus présentés précédemment, en modifiant ponctuellement des éléments de la tâche ou du déroulement de l'ÉTCSA, mais elle peut aussi impliquer de dévier du modèle type d'ÉTCSA. Cette situation s'est d'ailleurs posée pour Julien concernant le moment d'intégration de la tâche liée à l'ÉTCSA dans la séquence du cours : alors qu'au départ, il avait conçu le projet pour démarrer l'ÉTCSA dès le début de la session et de maintenir une alternance constante entre l'enseignement des notions théoriques et la réalisation de la tâche, il a par la suite modifié son approche afin de faire face aux imprévus.

[Dans la conception initiale], dès le départ, je voyais l'explication [des contenus associés au cours], au début de la session pour donner les outils. Mais assez rapidement, je rentrais dans [la tâche associée à l'ÉTCSA], effectivement, et là ça posait un peu plus de problèmes. Là c'est un peu plus tard, donc j'ai gardé l'ordre de présentation de la matière théorique, mais c'est un peu plus tard qu'on commence [les tâches associées à l'ÉTCSA], donc ils ont un peu plus le temps d'assimiler, de juste comprendre comment fonctionne... [...] aussi [si] un étudiant [...] quitte à la 3e semaine, qu'il puisse quitter sans que ça perturbe. (E.283-299)

Cette même nécessité de s'adapter au contexte d'administration rencontré l'a amené à revoir certaines dimensions de la tâche telle qu'il l'avait conçu initialement.

Ainsi, pour faire face aux imprévus, pour prendre en compte les capacités des étudiantes et des étudiants et leur rythme d'apprentissage ou en raison de contraintes liées au contexte académique et institutionnel, il faut parfois simplifier la réalité dans le cadre de l'ÉTCSA. C'est du moins le type de latitude que s'est donné Julien en regard des modèles en place. Par exemple, il a retiré une des fonctions incluses dans la première mouture de son ÉTCSA, car elle rendait la situation d'apprentissage trop complexe et créait de la confusion chez les étudiantes et étudiants. En faisant cela, Julien était très conscient de diminuer l'authenticité de la tâche et de son contexte de réalisation, mais il a fait ce choix au profit d'un fonctionnement plus adapté au contexte de la classe et au niveau des étudiantes et des étudiants.

Un résumé des éléments composant ces deux pôles est également présenté en annexe F (Tableau 3 – Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'administration de l'ÉTCSA).

1.3. Phase d'évaluation

La phase de l'évaluation a été définie comme «la phase durant laquelle l'évaluateur analyse l'ensemble de la production ou performance finale à partir de chaque indicateur afin de porter un jugement et sanctionner l'atteinte ou non des objectifs d'apprentissage et des standards visés et certifier de l'atteinte ou non de la/des compétence (s) associée (s) au cours.» (annexe C, p. 180)

Il ressort de l'analyse des données recueillies que la majorité des enjeux associés à cette phase sont reliés aux questions d'équité et de fiabilité. À ceux-ci, s'ajoutent quelques enjeux spécifiques à la première mise en œuvre, certaines considérations concernant le temps et les ressources, et quelques questionnements similaires à ceux engendrés par des évaluations plus traditionnelles. Un tableau résumant ces éléments du vécu de Julien concernant la phase l'évaluation est également

présenté en annexe H (Tableau 4 – Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'évaluation de l'ÉTCSA).

1.3.1. Les enjeux d'équité

Les propos du participant démontrent que les enjeux d'équité et de validité, ne sont pas cloisonnés à cette seule phase de la mise en œuvre d'une ÉTCSA, mais au contraire traversent l'ensemble des phases de la mise en œuvre. Toutefois ce thème a surtout été abordé dans la portion de l'entrevue concernant la phase d'évaluation, car c'est dans cette phase que Julien a été confronté de manière concrète aux résultats de ses décisions antérieures et leurs effets sur l'équité. Il a identifié quatre catégories d'enjeux liés à l'équité et à la fiabilité de l'évaluation dans le cadre de la mise en œuvre de son ÉTCSA, soient les enjeux d'équité liés à la non-standardisation des tâches réalisées par les étudiantes et étudiants, les impacts des comportements et des choix des étudiantes et étudiants sur l'équité et les enjeux liés au bagage socioculturel.

1.3.2. Enjeux liés à la non-standardisation des tâches réalisées par les étudiantes et étudiants

Le fait que le scénario d'ÉTCSA élaboré par Julien comporte plusieurs rôles, rattachés à des tâches différentes, entraîne des enjeux d'équité entre les étudiantes et les étudiants d'un même groupe. Dans le cas de l'ÉTCSA de Julien, cette différenciation implique l'utilisation de grilles d'évaluation présentant des variations selon les rôles, et il est très conscient des enjeux associés à cette situation. Ainsi, pour certaines évaluations, Julien utilise une même grille d'évaluation même si les consignes peuvent varier selon le rôle de l'étudiante ou de l'étudiant dans la tâche. Par contre, pour certaines tâches, les grilles utilisées sont différentes.

Surtout que dans ce cadre d'évaluation là, tous les étudiantes et les étudiants n'ont pas le même rôle, n'ont pas les mêmes tâches, s'assurer qu'en termes d'évaluation, dans on ramène ça dans un contexte d'apprentissage collégial, de cote R, etc., que ce soit relativement équivalent, ça, ça n'a pas été facile. Ça a été ça le plus gros défi. Toi si je

te demande de faire un travail de 3 pages là-dessus, [et que toi] je peux pas te demander de faire ce travail-là à cause de ton rôle, mais que je dois te demander de faire 2 travaux de 1 page et demi, là je simplifie beaucoup [mon explication] mais.... ça c'était la plus grande difficulté. (E.36-44)

Chez Julien, les enjeux liés à la non-standardisation des évaluations ont été une source de stress et ont suscité beaucoup de réflexion, à la fois sur le principe et sur la façon de maintenir une équité entre les étudiantes et les étudiants :

Ça fait peur au début.... c'est une façon d'évaluer quand même très différente, c'est un contexte d'apprentissage et d'évaluation très différent de ce qu'on fait traditionnellement.... Pis juste le fait que tous les étudiants ne sont pas évalués sur la même base, ça aussi ça fait en sorte que..... des profs peuvent même être contre ça en tant que tel, en disant que tout le monde doit être évalué sur les mêmes critères... comment je m'assure que c'est vraiment équitable ? (E.368-374)

Mais pour Julien, la qualité et l'authenticité des apprentissages réalisés à l'intérieur de son modèle d'ÉTCSA font amplement contrepoids à ces enjeux.

1.3.3. Impacts des comportements et des choix des étudiantes et étudiants sur l'équité

Un autre enjeu d'équité identifié par Julien concerne le fait que les comportements et les choix faits par les étudiantes et étudiants ont des répercussions sur leurs collègues. Certaines décisions quant à la façon de travailler peuvent aussi avoir des répercussions sur les résultats des étudiants : «les étudiants peuvent se parler entre eux et s'aider, mais ils ne le font pas beaucoup. En fait, certains groupes le font, j'ai des groupes [où] il y a de l'entraide, ils collaborent... il y a une partie que c'est du ressort des étudiants, d'utiliser les outils pour travailler ensemble.» (E.557-563) Ainsi, le vécu des étudiantes et des étudiants peut varier à l'intérieur d'un même groupe. Mais c'est la variabilité entre les groupes qui ressort davantage des propos de Julien. Il donne l'exemple du niveau de participation au sein d'un groupe, qui a des répercussions plus importantes sur le déroulement de la tâche et son contexte de réalisation dans le cas d'une ÉTCSA, comparativement à une approche plus traditionnelle de l'évaluation.

Si je demande de faire un travail de recherche de 5 pages, les étudiants vont me remettre des travaux plus ou moins bons, bien entendu, mais ça va être assez similaire.... [...] l'expérience du prof ou du groupe va être assez similaire. Tandis que là, si mes étudiants participent beaucoup, participent moins, ça change. Pour prendre un exemple très simple, si [dans un groupe] des étudiants commencent à citer beaucoup de sources. Ça va pousser les autres à [se dire] "C'est ça qu'il faut que je fasse" et ils vont citer beaucoup de sources. Et là, ça va améliorer collectivement [les résultats]. Si au contraire, personne ne le fait, ça démotive d'autres à le faire, et ils le font encore moins. Quand je fais un travail de 5 pages, ben je vais faire ce que je pense que je dois faire. Je peux en parler avec mes amis, mais je fais ce que je pense, et c'est ça. C'est plus individuel. [...] Pis c'est pour ça que c'est bien qu'ils s'aident, et que j'essaie de les pousser à ce qu'ils se parlent, qu'ils s'aident. Voyant ce que les autres font, et aussi voyant mes rétroactions [...] Collectivement, ils s'améliorent. C'est intéressant, ça. Parce que dans le fond, au niveau de l'équité à l'intérieur des groupes, il ne semble pas y avoir de problème, mais au niveau de l'équité, interprétée de façon large, entre les groupes, l'expérience d'un étudiant d'un groupe versus celui d'un autre groupe et les répercussions sur l'évaluation diffèrent... [...] C'est ça aussi qui est particulier, c'est que c'est public, dans le groupe bien entendu, dans le sens que tout le monde voit les travaux des autres. C'est la même chose si un étudiant me remet un bon [travail], ça va faire des débats intéressants. Un étudiant me remet un [travail moins bon], ça va être plus difficile d'intervenir [...] [...] mais le travail de chacun a plus d'impact sur les autres que quand ce sont des travaux individuels, même d'équipes. C'est que là, il n'y a pas juste une relation prof-élève, mais aussi une relation élève-élève. C'est ça l'ajout ici, par rapport à une relation plus traditionnelle. (E.719-761)

Ainsi, pour Julien, ce type d'enjeu concernant l'équité se pose surtout entre les groupes qui suivent un même cours. Pour tenter de réduire cet impact de la dynamique du cours, il insiste sur l'importance de maintenir le caractère individuel de l'évaluation et de la notation : « Même si c'est du travail collaboratif, ce sont toutes des évaluations individuelles. » (E.612-613)

Ces enjeux constatés par Julien l'ont amené à questionner les étudiantes et les étudiants pour recueillir leurs perceptions concernant l'équité des évaluations.

Les deux, trois premières sessions [où j'ai implanté cette évaluation], à la fin de la session, j'ai distribué un questionnaire à mes étudiants pour évaluer l'ensemble de l'œuvre. Et une des questions c'était "est-ce que vous trouvez que votre charge de travail était équivalente aux autres ou pas ?" Dans une vaste majorité, ils trouvaient que oui [c'était équivalent]. (E.678-682)

Cette perception semble confirmée du fait que Julien n'a pas fait face à des contestations officielles, ou des demandes de révision de note, depuis qu'il a mis en place l'ÉTCSA (E.686-699).

1.3.4. Les enjeux concernant le bagage socioculturel

En fin d'entrevue, Julien a été interrogé sur les enjeux liés au bagage socioculturel des étudiantes et étudiants, notamment concernant l'équité. Sur ce point, Julien ne perçoit pas de différence majeure entre la capacité des étudiantes et étudiants à bien réussir l'ÉTCSA :

Je dirais que [dans mon cégep] c'est très homogène. On est beaucoup dans du Québécois, de souche, blanc, avec des parents qui ont quand même un peu d'argent. Classe moyenne, classe moyenne-aisée. C'est le portrait-type. Pratiquement tous mes étudiants ont leur voiture à eux. Ça fait qu'il n'y a pas une grande diversité ici. [...] [Mais si j'avais à transposer mon projet dans un milieu où le bagage socioculturel est différent] a priori.... évidemment, c'est complètement subjectif ce que je dis, mais..... je penserais pas [qu'il y ait de problèmes]. [...] C'est plus au niveau de la motivation. Ceux qui aiment plus ça vont être plus motivés, et vont [s'impliquer davantage] et s'assurer de bien comprendre. C'est plus au niveau de la motivation que je vais voir quelque chose. (E.878-903)

Dans l'ensemble, bien que les enjeux d'équité aient été abordés principalement dans la phase de l'entrevue concernant l'évaluation, il ressort des propos de Julien que c'est surtout lors de la phase d'élaboration que se joue réellement l'équité. Pour lui, l'important pour prévenir les problèmes d'équité est de

bien penser ses consignes, bien réfléchir. Avoir le temps, de bien réfléchir. C'est ça que j'ai pu faire. Et jaser avec les collègues. J'ai des consignes que j'ai pu montrer à des collègues, qui m'ont dit [parfois] "il me semble qu'il manque quelque chose ici, il me semble que cette tâche-là est plus facile que celle-là". "Tu as peut-être raison, peut-être que si je rajoute cette consigne, cette exigence-là".... "ah oui" (E.702-707)

Ici, Julien rejoint les éléments abordés dans la phase d'élaboration de l'ÉTCSA, particulièrement concernant les conditions favorables.

1.3.5. Les enjeux liés à la première mise en œuvre

À l'instar de la phase d'administration, Julien mentionne que la première expérience de mise en œuvre présente des difficultés plus importantes que les subséquentes. Cet écart entre la première mise en œuvre et les subséquentes se manifeste notamment en termes de temps nécessaire à la phase d'évaluation et de notation. Cela ressort d'ailleurs très clairement dans le questionnaire qu'il a complété. En effet, il se déclare «fortement en accord» avec l'affirmation «la mise en place d'une ÉTCSA a occasionné une augmentation du nombre d'heures consacrées à l'évaluation sommative et à la notation des épreuves terminales de cours lors de ma première expérience d'implantation» (Q13g-i). Toutefois, pour les mises en œuvre subséquentes, la phase d'évaluation ne semble pas demander significativement plus de temps dans cette approche que dans une approche plus traditionnelle de l'évaluation : en effet, Julien se déclare «assez en désaccord» avec l'affirmation selon laquelle «la mise en place d'une ÉTCSA a occasionné une augmentation du nombre d'heures consacrées à l'évaluation sommative et à la notation des épreuves terminales de cours lors des expériences d'implantation subséquentes» (Q13g-ii). Il s'agit d'ailleurs de l'écart le plus important concernant les réponses aux questions où Julien devait évaluer les répercussions de l'implantation de son projet d'ÉTCSA de manière différenciée entre la première expérience de mise en œuvre et les subséquentes.

En plus de la question du temps, Julien souligne aussi qu'il y a une amélioration de la qualité et de la rigueur de son évaluation après plusieurs répétitions de l'ÉTCSA :

Plus ça va mieux c'est. L'inverse serait très triste. C'est sûr qu'au début, c'était.... pas brouillon, mais plus expérimental. Tandis que là, moi-même, comme je disais tantôt, je sais plus où je m'en vais. Je sais telle semaine qu'est-ce que je suis censé faire, et ainsi de suite, et je suis plus capable d'encadrer mes étudiants. Je sais qu'il va se passer ça, qu'il va se passer ça, que cette consigne-là, c'est ça que je veux. Au début, tu le sais, mais tu n'as pas pensé à tout parce qu'on est humain. Plus d'expérience, ça va mieux. (E.708-716)

1.3.6. Les enjeux liés au temps et aux ressources

Bien que Julien ne perçoive pas les enjeux en temps et en ressources de la phase d'évaluation comme aussi aigus que pour la phase d'élaboration, certaines considérations relatives au temps et aux ressources associées à la phase d'évaluation ont tout de même été soulevées. Toutefois, elle concerne davantage l'organisation logistique et temporelle de celle-ci et non la lourdeur, comme c'est le cas pour les phases d'élaboration et d'administration. Pour Julien, contrairement à une approche plus traditionnelle, dans le cadre de son projet d'ETCSA, l'évaluation est plus fréquente et répartie sur une portion plus grande de la session, mais à plus petite échelle à chaque fois :

[Voici] comment je procède en termes d'évaluation [...] En cours de session, je lis, je fonctionne dans une logique formative, je lis les différentes interventions, en diagonale, rapide, pour savoir qu'est-ce qu'ils disent, et là, je corrige le tir. [Sous forme d']un appel au groupe, 'attention à ça, ça, ça''. Si je vois qu'un étudiant fait souvent la même erreur, je vais le contacter personnellement. [...] Mais je n'attribue pas de note à ce moment-là. Comme mes évaluations, de toute façon, sont globales, j'attends à la fin. À la fin, [pour chaque étudiant] je vais voir l'ensemble des interventions. Je les ai déjà lues une première fois, donc ça me permet d'aller plus vite, et là, j'attribue une note. Ça fait en sorte que je suis un peu toujours en évaluation, mais c'est une évaluation un peu plus légère. [...] Ça fait, qu'au final, je trouve que ça me prend moins de temps à corriger [au moment de l'évaluation finale] (E.457-476)

Finalement, mentionnons que sur le plan des difficultés liées à la phase d'évaluation, Julien mentionne qu'il a également rencontré des difficultés similaires à celles rencontrées dans une approche traditionnelle en matière de validité de l'évaluation :

quels critères (utiliser) ? Comment formuler ça ? Mais bon, ça c'est dans n'importe quelle évaluation, de toute façon, c'est l'expérience..... Et certaines consignes étaient plus ou moins claires finalement. Il y a plusieurs consignes que j'ai reprécisées par la suite, pour lesquelles j'ai rajouté de l'information, j'ai rajouté des exemples.... (E.627-631)

Selon Julien, ces enjeux plus classiques ne sont pas plus importants dans une approche d'ÉTCSA que dans une approche traditionnelle, mais elles ne sont pas non plus évacuées, et doivent être abordées.

Un tableau résumant ces éléments du vécu de Julien concernant la phase d'évaluation est présenté en annexe H (Tableau 4 – Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'évaluation de l'ÉTCSA).

2. ANALYSE ET INTERPRÉTATION DES RÉSULTATS

Afin de compléter l'étude des conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans des cours de première année du programme d'études préuniversitaires en sciences humaines au collégial, il est pertinent de mettre en relation les données recueillies dans la présente recherche avec les éléments recensés dans la littérature concernant l'ÉSA. Pour ce faire, une comparaison entre les données recueillies auprès de Julien et les éléments de la littérature a été effectuée pour les quatre grandes catégories d'enjeux associés à la mise en œuvre d'une ÉSA, présentés

dans le chapitre 2. Les aspects attendus et inattendus des résultats obtenus seront également présentés.

2.1. La qualité et la rigueur du processus évaluatif

Rappelons que les enjeux en matière de qualité et de rigueur du processus évaluatif recensés dans la littérature sur l'évaluation en situation authentique s'articulent principalement autour des problèmes de fiabilité, de validité et d'équité. Voyons maintenant comment les éléments traités par la littérature à ce propos peuvent être mis en relation avec les résultats présentés précédemment.

Au niveau des enjeux de fiabilité, ils sont abordés à la fois dans la littérature et dans les données recueillies, dans les deux cas comme effet de la grande hétérogénéité du produit de l'ÉSA. Toutefois, si la cause identifiée est similaire, la teneur des préoccupations est différente. En effet, le type d'enjeux reliés à la fiabilité qui préoccupe les autrices et les auteurs est différent de celui qui est présenté par Julien. En effet, des auteurs centraux de la littérature sur l'ÉSA, tel Tanner (2001), Hoepfl (2000), Wiggins (2011) et Montgomery (2002), ont abordé la question de la fiabilité, principalement dans la perspective de la concordance des jugements entre les évaluateurs: la principale préoccupation recensée en la matière dans les écrits de ces auteurs concerne les écarts de jugement possibles entre différents évaluateurs. Conséquemment, les solutions proposées par ceux-ci concernaient les moyens d'atteindre une plus grande homogénéité entre les jugements portés sur le produit d'une ÉSA.

Or, dans le cas des données recueillies, les principaux questionnements concernent davantage la fiabilité des évaluations selon les différents rôles et tâches. Ainsi, il rejoint davantage les préoccupations d'auteurs comme Perkins et Salomon (1989, dans Scallon, 2004) et de Linn et Burton (1994, dans Scallon, 2004) concernant les enjeux de fiabilité relevant de «d'interaction personne x tâche» (Scallon, 2004),

traité dans la littérature sur l'évaluation des performances en générale, et non de la littérature spécifique à l'évaluation en situation authentique, ce qui est un aspect inattendu des données recueillies. Toutefois, le constat de toutes ces autrices et tous ces auteurs est similaire, peu importe le type de problème de fiabilité : la majorité accepte le sacrifice d'une certaine part de fiabilité au profit de la validité des évaluations, qui est supérieure aux évaluations plus traditionnelles ((Frey *et al.*, 2012; Montgomery, 2002; Newmann, Brandt et Wiggins, 1998; Palm, 2008; Riley et Slater Stern, 1998; Scallon, 2004; Wiggins, 1990, 1993a). C'est aussi le cas des données recueillies dans la présente recherche, où le compromis est perçu comme un sacrifice acceptable face à cette problématique. Toutefois, s'il y a convergence dans la solution, les justifications invoquées dans les données de la présente recherche diffèrent de celles recensées dans la littérature, car l'acceptation d'une certaine marge d'erreur en matière de fiabilité repose davantage sur l'importance de maintenir une équité entre les étudiantes et étudiants plus que sur la base de la validité : ainsi, dans nos résultats, c'est dans la mesure où il y a une perception d'équité validée par les pairs et par les étudiantes et étudiants, que les risques associés à l'hétérogénéité des produits et des tâches, et donc de potentiels écarts quant à la fiabilité, sont acceptables. Cette dimension constitue un aspect inattendu des résultats obtenus dans la présente recherche.

Dans la littérature, la fiabilité des évaluations est aussi abordée à travers la question des biais induits par le contexte de réalisation (Hoepfl, 2000; Messick, 1996, Yen, 1993, dans Hoepfl, 2000) : ici, c'est la fiabilité des résultats et les problèmes de généralisation de l'évaluation à partir d'une seule situation de performance, dont on peut difficilement garantir qu'il est assurément représentatif de l'apprentissage global réalisé, qui sont questionnés. Or, dans une approche ÉSA, en raison du temps et des ressources nécessaires à la réalisation d'une tâche évaluative, il est difficile de multiplier les performances et productions afin de s'assurer de cette forme de fiabilité de l'évaluation.

On retrouve également cette préoccupation dans les données recueillies, mais elle est formulée non pas en tant qu'enjeu de fiabilité, mais en tant qu'enjeu d'équité entre les étudiantes et les étudiants, notamment en raison du contexte de comparaison à travers des outils comme la Cote R. Ainsi, s'il y a convergence quant aux constats entre nos données et la littérature, soit la préoccupation découlant du caractère contextualisé de la tâche et de l'hétérogénéité des situations et des productions, cela n'est pas associé à la même dimension de l'évaluation (problème de fiabilité versus problème d'équité). Ceci est une dimension inattendue des résultats de la présente recherche.

Les solutions et conditions favorables face à ces enjeux identifiés dans la littérature et dans les données de la présente recherche sont d'ailleurs différentes : dans le premier cas, c'est l'importance de la formation des évaluateurs et de la multiplication des situations évaluées qui est promue, alors que dans les données recueillies, la priorité est de valider l'équivalence des tâches auprès de pairs et d'évaluer la perception des étudiantes et étudiants. Ces approches, bien que différentes, ne sont pas incompatibles, et dénotent avant tout une position différente dans le questionnement : alors que les autrices et auteurs envisagent des solutions demandant des interventions plus systémiques en raison des ressources nécessaires pour les mettre en œuvre, les données de la présente recherche se tournent vers des solutions d'envergure plus modeste, qui peuvent être mises en œuvre à plus brève échéance par un enseignant ou une enseignante.

Finalement, une autre différence entre la littérature et les propos recueillis dans la présente recherche est que, dans la littérature, la question de la fiabilité est principalement abordée en lien avec la phase d'évaluation, alors que dans les données recueillies, elle concerne aussi la phase d'élaboration. Ceci est une des dimensions les plus surprenantes de nos données, qui n'avait pas du tout été anticipée.

2.2. Exigences en temps et en ressources

La question des ressources est présente dans la littérature, tout comme dans le matériel recueilli dans la présente recherche, et il y a, dans l'ensemble, concordance entre les éléments.

Outre Wiggins (1993a), pour qui l'argument de la nécessité de ressources supplémentaire est en fait une forme de résistance au changement, l'ensemble de la littérature ayant abordé cette question s'accorde sur le fait que la mise en place d'une ÉSA demande plus de ressources ((Ewing, 1998; Hardy, 1996; Khattri *et al.*, 1998; Maeroff, 1991; Monk, 1996; Montgomery, 2002; Tanner, 2001; Taylor, 1994, Rothman, 1995, dans Hoepfl, 2000; Popham, 1993), et que ce besoin en ressources se fait principalement sentir lors de la phase d'élaboration de l'ÉSA (Hardy, 1996; Hoepfl, 2000; Maeroff, 1991; Monk, 1996). Les données recueillies concordent parfaitement avec ces éléments de la littérature: en effet, le fait de disposer d'une libération (dégrèvement) des tâches d'enseignement pour pouvoir concevoir un projet d'ÉTCSA comme le facteur le plus important dans la capacité de l'enseignante ou de l'enseignant à mener à bien ce projet. Autre élément de consensus entre la littérature et les données recueillies : les ressources nécessaires à la conception d'un canevas d'ÉSA sont relativement incompressibles, et ne varient pas selon le nombre d'étudiantes et étudiants auquel il sera administré.

Sur le plan des ressources associées à la phase d'administration, il y a une parenté claire entre les constats se dégageant de la littérature et ceux relevant des données recueillies sur le fait que l'administration d'une ÉSA demande plus de soutien et d'accompagnement pédagogique (Hardy, 1996; Hoepfl, 2000; Khattri *et al.*, 1998; Monk, 1996). Dans les deux cas, ce sont les besoins en termes de rétroaction fréquente et de soutien constant à l'apprentissage qui explique cela. Une nuance importante est toutefois présente dans le matériel recueilli dans la présente recherche : il y a une distinction claire entre la première implantation, beaucoup plus exigeante, et les mises

en œuvre subséquentes. Or, cette distinction entre la première implantation et les mises en œuvre subséquentes n'est pas présente dans la littérature consultée. Toutefois, cet aspect avait été anticipé dès les premières étapes de la présente démarche de recherche, au point où des questions portant spécifiquement sur les différences entre la première mise en œuvre et les suivantes avaient été intégrées au questionnaire et à l'entrevue. Il ne s'agit donc pas d'un résultat inattendu, bien que cette distinction ne se retrouvait pas dans la littérature consultée.

Autre écart entre la littérature et les données de la présente recherche: dans la littérature, le soutien nécessaire lors de la phase d'administration implique d'ajouter des ressources professionnelles ou de limiter le nombre d'étudiantes et étudiants par groupe. Outre le cas de la première mise en œuvre, les données recueillies ne sont pas aussi tranchées - la variation du nombre d'étudiantes et étudiants par classe n'ayant pas été identifiée comme ayant plus d'impact sur sa capacité à mener à bien la phase d'administration de son ÉTCSA que dans une approche traditionnelle de l'évaluation – que dans la littérature. Toutefois, il importe de garder en tête que, dans le cas à l'étude dans la présente recherche, cette variation est de quelques étudiantes et étudiants par session, alors que dans la littérature, on parle de variations beaucoup plus considérables, la majorité de ces autrices et ces auteurs s'étant attardé à des expériences menées à grande échelle (pour l'ensemble d'un secteur ou d'un État) (Hardy, 1996; Hoepfl, 2000; Khattri *et al.*, 1998; Monk, 1996). Ainsi, bien que les propos ne concordent pas parfaitement sur ce point, ils ne sont pas nécessairement opposés pour autant.

C'est sur le plan des ressources associées à la phase d'évaluation que l'écart entre la littérature et l'expérience du participant diffère le plus. En effet, pour plusieurs autrices et auteurs (Hardy, 1996; Hoepfl, 2000; Kerka, 1993; Khattri *et al.*, 1998; Maeroff, 1991; Monk, 1996; Popham, 1993), la phase d'évaluation nécessite substantiellement plus de temps dans une approche en situation authentique que dans une approche psychoéducative de l'évaluation. Or, dans les données recueillies, outre

la première expérience de mise en œuvre d'une nouvelle ÉTCSA, le temps nécessaire à la réalisation de la phase d'évaluation des productions des étudiantes et étudiants n'était pas supérieur à l'approche utilisée antérieurement. Ce résultat était très inattendu, et a nécessité une analyse plus approfondie afin de comprendre cet écart avec la littérature. Une piste d'explication peut venir du fait que le cas à l'étude, l'ÉTCSA était organisée autour de groupe collaboratif, bien que l'évaluation était individuelle, ce qui est une des approches proposées par Hardy (1996) et Wiggins (1990, 2011) pour diminuer le temps nécessaire à l'évaluation des productions résultant d'une ÉSA. Cela est une piste d'explication concernant le fait qu'aucun alourdissement de la tâche lors de la phase d'évaluation n'a été documenté dans les données recueillies. Une autre piste est le fait que dans la littérature, les ressources nécessaires à l'évaluation dans une approche authentique est souvent comparée à celles mobilisées dans une approche psychoéducative, alors que dans le cas à l'étude dans la présente recherche, l'expérience documentée est comparée avec le vécu antérieur qui s'inscrivait déjà dans une approche d'évaluation des performances, qui est en soit une approche demandant plus de temps et de ressources que l'approche psychoéducative au niveau de la phase d'évaluation. Ainsi, dans les données recueillies, la base de comparaison n'est pas la même que dans la littérature, ce qui implique que l'écart entre les propos n'est pas nécessairement incompatible. Demeure tout de même qu'il y a un écart entre les constats des études comme celles des projets pilotes américains et britanniques dont font état des auteurs comme Hardy (1996), Hoepfl (2000), Maeroff (1991) et Popham, 1993), dont certains désignent même les coûts additionnels associés à la phase d'évaluation d'ÉSA comme étant un obstacle important - voire prohibitif - à l'implantation de cette approche évaluative à large échelle, et ceux de la présente étude, dont les données ne démontrent pas que cette charge est significative dans le facteur de faisabilité de l'implantation d'une ÉSA.

Finalement, une autre dimension relative aux ressources concerne celles nécessaires à la mise en place d'un contexte adapté à la tâche authentique (Hardy, 1996; Hoepfl, 2000). Ces coûts sont très variables, car ils dépendent du type de contexte

nécessaire à la tâche et les informations présentes dans la littérature concernent des expériences spécifiques, donc peu généralisables. Dans le cas à l'étude dans la présente recherche, le choix de procéder par simulation en ligne, plutôt que par la mise en place d'un contexte parfaitement authentique, a permis de contourner cet obstacle, qui aurait été insurmontable, advenant le choix d'une tâche parfaitement authentique dans son contexte. Ainsi, dans les données recueillies, ce sont davantage les ressources humaines, notamment sur le plan du temps disponible, qui sont au cœur du discours. Sur ce point, en raison du haut degré de variabilité de ce facteur, il n'est pas possible de faire une comparaison significative entre la littérature et la situation du participant, ni même entre les différents textes de la littérature, d'ailleurs.

2.3. Le contexte et les contraintes institutionnelles

Dans la littérature répertoriée à propos de l'ÉSA en contexte académique postsecondaire, peu d'autrices et d'auteurs abordent explicitement les contraintes institutionnelles et leurs impacts sur la mise en place d'ÉSA, mais ceux qui le font traitent principalement des difficultés associées aux courants de standardisation des modalités d'évaluation, ayant encouragé la mise en place de mécanismes dont les normes sont parfois si précises qu'elles n'offrent pas la latitude nécessaire à la conception d'ÉSA, particulièrement lorsqu'elles impliquent une part importante de travail collaboratif (Clarke *et al.*, 2000; Herrington et Herrington, 2006). Or, cette tendance est très présente dans le réseau collégial québécois, avec l'arrivée des exigences pour les institutions de se doter d'une politique d'évaluation des apprentissages très détaillée (Gouvernement du Québec, 2013). Toutefois, les données recueillies concernant le cas à l'étude dans la présente recherche ne démontrent pas d'entraves majeures relevant de cette tendance. En effet, dans le cas à l'étude, les démarches qui ont été rapportées se sont bien déroulées et étaient relativement simples et rapides. Toutefois, la présence d'une ouverture au sein de l'institution face à une certaine variabilité des tâches pour une même évaluation a été identifiée comme centrale dans la possibilité de mettre sur pied un projet d'ÉTCSA : plus encore, cette

flexibilité pédagogique est identifiée comme nécessaire à la mise en place d'une ÉTCSA. Ainsi, les conditions identifiées comme favorables dans la littérature – donner préséance à une approche centrée sur l'apport du contexte d'apprentissage plutôt que sur une perspective individuelle de l'apprentissage et de l'évaluation – étaient réunies dans le cas à l'étude dans la présente recherche, et cela a bel et bien été une condition favorable identifiée dans les données recueillies : ici, il y a concordance entre les éléments de la littérature et les résultats de la présente recherche.

Dans la présente recherche, deux éléments relevant du contexte institutionnel non documenté dans la littérature ont été recensés. La première est une difficulté relevant spécifiquement du contexte institutionnel des études collégiales : le format des cours. Dans le cas à l'étude, le fait que le devis du cours prévoit 3 heures en classe par semaine et 3 heures de travail à la maison restreignait la possibilité de réaliser en classe certaines dimensions de la tâche. Cela a eu pour effet de déplacer une partie des activités sur une plateforme en ligne, diminuant du coup le caractère authentique de celle-ci sur le plan du contexte de réalisation (dans le contexte extrascolaire, le type d'interaction de la tâche est plutôt réalisé en présence des intervenants et non de manière virtuelle).

La seconde difficulté relève du système d'admission universitaire, qui met, en quelque sorte, les étudiantes et étudiants en compétition entre eux, ce qui peut rendre délicat le fait d'évaluer une même compétence de manière non standardisée au sein d'un même groupe, particulièrement dans un programme préuniversitaire comme le programme de sciences humaines. Bien que dans le cas à l'étude, il n'y ait pas eu de plainte de la part des étudiantes et étudiants à ce propos, l'adoption d'une approche moins standardisée a tout de même été source de longues réflexions lors de la phase de conception en raison de cet aspect du contexte institutionnel.

Ces deux derniers éléments ne se retrouvent pas dans la littérature consultée, situation qui pourrait être en partie due au fait que certaines dimensions de ces

difficultés découlent du système collégial québécois, alors que la littérature sur l'ÉSA est principalement américaine. Toutefois, ceci n'est pas une découverte inattendue, car elle est tout à fait cohérente avec le contexte québécois, et qu'il était attendu de retrouver des éléments qui ne seraient pas couverts dans les études traitant des réalités non québécoises.

2.4. Les problèmes de nature pédagogique

Sur le plan des problèmes de nature didactique et pédagogique, les données recueillies rejoignent les deux grandes catégories d'enjeux abordés dans la littérature : soit les accros à l'authenticité de la tâche conçue et les impacts du manque d'expérience des apprenants en matière d'ÉSA. Ainsi, sur ce thème, les résultats obtenus ne présentent pas d'aspects non attendus.

Certaines autrices et certains auteurs (Cumming et Maxwell, 1999; Herrington et Herrington, 2006; Newmann et Archbald, 1992; Wiggins, 1993*a*, 1993*b*) ont constaté que les scénarios d'ÉSA conçus par les praticiennes et praticiens ne présentent pas toujours un degré d'authenticité optimal. Cette situation peut parfois être causée par un focus disproportionné sur des éléments du contexte plutôt que sur les connaissances et habiletés à maîtriser ou une mauvaise connaissance du concept d'ÉSA. Dans le même sens, certains scénarios peuvent n'être que du camouflage visant à donner une apparence d'authenticité à une tâche en ne lui ajoutant qu'une contextualisation sommaire (Cumming et Maxwell, 1999). Dans de tels cas, l'ÉSA produite s'en trouve affecté du fait que ces biais introduisent dans l'évaluation des enjeux ne relevant pas réellement de la situation authentique ciblée par l'apprentissage.

On peut ici établir un parallèle entre le choix du participant à notre étude de se rabattre sur une plateforme numérique pour la réalisation d'une portion de la tâche, du fait que le contexte institutionnel n'offre pas suffisamment d'heures en classe pour procéder dans un contexte le plus authentique possible. Ici, un enjeu relevant en

premier lieu du contexte institutionnel engendre aussi un problème de nature pédagogique concernant le degré d'authenticité du contexte de réalisation de la tâche. Et bien que la cause soit différente dans l'étude de cas de celles identifiées dans la littérature (qui relève plutôt d'erreur de conception ou de compréhension du concept, ce qui n'est pas le cas chez le participant à cette étude, qui a opté consciemment pour ce scénario altérant l'authenticité de la tâche du fait que le scénario idéal était impossible dans le cadre qui lui était imparti), le résultat est le même, soit une altération du degré d'authenticité global de l'ÉTCSA. Et par le fait même, l'introduction d'un biais de validité dans la tâche et l'évaluation en résultant, tel qu'identifié par la littérature.

Concernant les solutions identifiées pour pallier ce problème, du fait que dans le cas à l'étude dans la présente recherche, la situation découle d'un élément contextuel sur lequel l'enseignant interviewé n'a pas d'emprise, la piste de solution identifiée dans la littérature – soit une conception soignée basée sur les réelles caractéristiques de l'ÉSA afin que la tâche conçue et les critères d'évaluation reliés soit équilibrés et réellement représentatifs du domaine d'expertise concerné par l'apprentissage – ne permet pas de remédier à cet enjeu.

Finalement, en ce qui concerne les impacts du manque d'expérience des étudiantes et étudiants en matière d'ÉSA, faute d'avoir été souvent exposés à cette formule durant leur parcours scolaire, on retrouve bel et bien des traces de cet enjeu dans les données recueillies. Et tout comme les autrices et les auteurs consultés à ce propos (Liftig, *et al.*, 1992, dans Montgomery, 2002), les données recueillies démontrent que cela engendre de l'insécurité et un grand besoin d'accompagnement chez les étudiants et étudiantes. En effet, durant la phase d'administration de l'ÉTCSA, une partie des difficultés recensées relevaient du fait que l'enseignant avait surestimé certaines connaissances et habiletés des étudiantes et étudiants nécessaires à une appropriation rapide de la tâche demandée et de son contexte. Il faut toutefois mentionner qu'une partie de ces difficultés relevaient davantage du manque de connaissance du contexte institutionnel propre au niveau collégial ou de connaissances

plus limitées au niveau technologique. Toutefois, il ressort clairement des données qu'il y a un grand besoin d'accompagnement de la part des étudiantes et étudiants, pour des motifs similaires à ceux mentionnés dans la littérature, soit la grande latitude dont disposent les étudiantes et étudiants, le caractère appliqué et inédit de la tâche et le niveau d'habileté intellectuel (Liftig, Liftig et Eaker 1992, dans Montgomery, 2002). Les résultats présentés précédemment font d'ailleurs état de ces difficultés, associés à cet important besoin d'accompagnement.

Sur le plan des pistes de solution soulevées dans la littérature pour faire face à cet enjeu pédagogique, les deux éléments avancés (augmentation de l'exposition des étudiantes et étudiants à des ÉSA et augmentation des ressources en phase d'implantation) n'étaient pas à la portée du participant à cette recherche. Celui-ci a donc plutôt opté pour une diminution de certaines dimensions de la tâche, afin de circonscrire celle-ci et diminuer les besoins de soutien de la part des étudiantes et étudiants. Il est toutefois intéressant de constater que l'apport de ressources additionnelles durant la phase d'implantation comme condition favorable à la mise en œuvre d'une ÉTCSA est aussi validée par les données recueillies, qui identifient la présence d'un stagiaire d'expérience présent en classe avec lui lors de la première mise en œuvre de son projet d'ÉTCSA comme un facteur de la réussite de cette expérience : toutefois, la possibilité de disposer de ressources additionnelles sur une base récurrente étant difficile dans le contexte collégial actuel, la diminution de l'ampleur et de la complexité de la tâche s'est avérée une avenue plus pérenne dans le contexte d'enseignement habituel documenté dans la présente recherche. Ainsi, bien que les conditions favorables et les pistes d'action identifiées dans la littérature ne soient pas les mêmes que celles documentées dans le cas à l'étude ici, car relevant d'un niveau décisionnel différent (institutionnel versus individuel), elles ne sont toutefois pas incompatibles.

CONCLUSION

En conclusion, il importe de faire un retour sur la problématique, le cadre de référence, la méthodologie et les résultats principaux de cet essai, pour ensuite présenter les limites de la démarche et proposer des pistes de réflexion pour des démarches d'approfondissement futures.

1. LES ÉLÉMENTS CENTRAUX DE CET ESSAI

Si l'approche par compétences, incarnée par la réforme de 1993 intitulée le Renouveau de l'enseignement collégial, a engendré un mouvement de transformation des pratiques pédagogiques chez les enseignantes et enseignants du collégial, plusieurs études recensent une transformation plus lente lorsqu'il est question d'évaluation sommative et certificative. Cet essai s'est intéressé aux facteurs qui favorisent et défavorisent la mise en œuvre d'évaluation terminale de cours en situation authentique (ÉTCSA), une approche compatible avec les orientations de la réforme et de l'approche par compétences, et tente de répondre à la question suivante : quelles sont les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'évaluation en situation authentique dans des cours de première année du programme préuniversitaire en sciences humaines au collégial ?

Pour y répondre, l'objectif de recherche principal suivant a été retenu : documenter les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique (ÉSA) dans des cours de première année du programme d'études préuniversitaires en sciences humaines au collégial.

À cet objectif général se greffent trois objectifs spécifiques : dans le cadre de la mise en œuvre de pratiques évaluatives s'appuyant sur les principes d'une évaluation en situation authentique par un enseignant ou une enseignante ayant expérimenté la mise en œuvre d'une telle évaluation dans le cadre d'un cours de première année du programme d'études préuniversitaires en de sciences humaines, 1) identifier des conditions et actions perçues comme favorables, 2) identifier les enjeux et difficultés rencontrées, et 3) analyser les enjeux et difficultés rencontrées et les conditions et actions perçues comme favorables.

Au niveau du cadre de référence qui a guidé la présente démarche, la genèse du concept d'évaluation en situation authentique, et une recension des écrits à ce propos de ce concept, ont d'abord été présentées. Ensuite, les deux principaux modèles d'Évaluation en situation authentique (ÉSA) les plus influents, celui de Archbald et Newmann et celui de Wiggins, ont été présentés et comparés. Finalement, les problèmes associés à la mise en œuvre de pratiques évaluatives conformes à cette approche ont été recensés et présentés.

Afin d'atteindre l'objectif de recherche poursuivi, une étude de cas unique a été réalisée auprès d'un enseignant, surnommé Julien, d'une discipline du programme de sciences humaines ayant expérimenté la conception et la mise en œuvre d'une évaluation terminale de cours en situation authentique dans le cadre d'un cours de première année de ce programme. Celui-ci a complété un questionnaire, participé à une entrevue semi-dirigée et fourni un dossier d'artefact, lesquels ont constitué les données analysées à partir d'une approche qualitative.

Les résultats de la collecte de données ont été présentés par phase de mise en œuvre d'une ÉTCSA, soit l'élaboration, l'administration et l'évaluation. Au niveau de la phase de conception, Julien identifie quatre enjeux majeurs : le temps, l'interaction avec des collègues, la disponibilité de ressources spécialisées et la latitude sur le plan institutionnel.

Parmi ces éléments, avoir du temps réservé à l'élaboration d'un projet d'ÉTCSA est le facteur le plus déterminant, selon Julien. Il identifie le fait d'avoir bénéficié d'une libération pour projet spécial, lui permettant de consacrer une partie importante de son temps de travail à la conception de son projet, comme une condition *sine qua non* à la réussite de celui-ci.

Du côté de l'interaction avec les collègues, bien que Julien était le principal architecte du projet et travaillait principalement seul, il a consulté ponctuellement des collègues et des ressources spécialisées. Il identifie cette possibilité de collaboration comme un facteur favorable à la réussite de la phase d'élaboration d'une ÉTCSA, car elle a permis de valider la viabilité des certaines dimensions du projet, la convivialité de certaines consignes et documents et la faisabilité logistique de certaines dimensions du projet, et de faire des modifications au besoin. Ces interactions lui ont aussi permis d'assurer l'équité des évaluations sur le plan de leur conception. Il identifie d'ailleurs cette possibilité de consulter les collègues comme le plus important rempart en matière d'équité des évaluations, au niveau de ce qui peut être contrôlé lors de la conception de l'ÉTCSA.

En plus des collègues, Julien a eu recours à l'aide de ressources spécialisées au sein de son institution. Il identifie cette aide comme importante dans le succès de son projet, notamment en raison de volet informatique de celui-ci. Selon lui, bien que le type de ressources spécialisées nécessaire puisse varier selon la nature du projet, c'est un facteur à ne pas négliger pour le succès d'un projet d'ÉTCSA.

Finalement, Julien, le participant à cette étude de cas unique, a souligné l'importance de disposer d'une latitude institutionnelle lorsque l'on entreprend de concevoir une ÉTCSA, notamment sur le plan du respect de la diversité des pratiques pédagogiques et la tolérance vis-à-vis une certaine variabilité des tâches réalisées par les étudiantes et étudiants. Ce fut le cas dans son institution d'enseignement, et il

considère que cela a été important dans sa décision d'entreprendre ce projet et sa capacité de le mener à bien.

Au niveau de la phase d'administration, on constate que les enjeux se divisent en deux temps, soient la première mise en œuvre et les mises en œuvre subséquentes. Selon Julien, la première mise en œuvre est particulièrement exigeante et présente les enjeux spécifiques suivants : les difficultés et besoins non anticipés des étudiantes et étudiants et la maîtrise imparfaite de la démarche et du matériel de la part de l'enseignante ou de l'enseignant. Le premier enjeu concerne le fait que malgré toute la préparation effectuée lors de la phase d'élaboration de l'ÉTCSA, l'enseignante ou l'enseignant ne peut pas avoir tout prévu, et des éléments non anticipés surviendront inévitablement, notamment sur le plan des besoins et difficultés des étudiantes et étudiants. À cela s'ajoute la maîtrise imparfaite de la démarche et du matériel par l'enseignante ou l'enseignant, inhérente à l'utilisation de nouveau matériel pédagogique. Afin de faire face à ces enjeux spécifiques à la première mise en œuvre, Julien insiste sur l'importance de se réserver du temps pour réagir aux imprévus et, lorsque c'est possible, de disposer de ressources humaines supplémentaires en classe (dans son cas, la présence d'un stagiaire a beaucoup aidé).

Au niveau des enjeux associés à la phase d'administration lors des mises en œuvre subséquentes, Julien rapporte certains enjeux récurrents ou contextuels, parmi lesquels un besoin accru de soutien et d'accompagnement de la part des étudiantes et étudiants, la variabilité du contexte pédagogique et les imprévus et l'importance des ressources humaines et matérielles suffisantes. En effet, il constate que la mise en œuvre d'une ÉTCSA demande plus de soutien et d'accompagnement des étudiantes et étudiants, comparativement à une épreuve terminale de cours plus traditionnelle. Selon lui, cela est dû en partie à la nature même de ce type d'évaluation, mais aussi au fait que plusieurs étudiantes et étudiants n'ont peu ou jamais été exposés à ce type d'approche dans leur parcours scolaire : il faut donc les accompagner davantage au niveau de la démarche. La variabilité du contexte pédagogique et les imprévus (par

exemple, les abandons en cours de session, sans avertissement) qui y sont associés ont également causé certaines difficultés à Julien. Pour pallier ces difficultés, Julien insiste sur l'important pour l'enseignante ou l'enseignant de disposer d'une latitude individuelle pour faire face aux imprévus de la phase d'administration. Il suggère aussi de démarrer les premières étapes de l'ÉTCSA seulement après quelques semaines de cours, une fois que la classe est stabilisée. Finalement, un enjeu de la phase d'administration d'une ÉTCSA est l'accès aux ressources matérielles et humaines adéquates. Dans son cas, c'est l'accès aux laboratoires informatiques et à un soutien spécialisé qui a été important, mais il souligne que la nature des ressources nécessaires peut varier selon le type d'ÉTCSA.

Finalement, concernant la phase d'évaluation, trois enjeux sont identifiés. Le premier concerne le principe d'équité. Dans le cas précis de son ÉTCSA, la non-standardisation des rôles entraîne des enjeux d'équité entre les étudiantes et les étudiants d'un même groupe lors de l'évaluation. De plus, il constate que les comportements et les choix des étudiantes et étudiants durant le déroulement de l'ÉTCSA ont un impact sur les autres étudiantes et étudiants, ce qui engendre des enjeux d'équité entre les groupes d'un même cours. Selon lui, la principale mesure à prendre pour réduire les enjeux d'équité est de porter une attention particulière lors de la phase d'élaboration (notamment en consultant des collègues) afin de prévenir ce qui peut être anticipé. Toutefois, selon lui, une part de ces enjeux demeure irréductible, mais il croit que les bénéfices de l'approche dépassent ces inconvénients.

Un autre constat associé à la phase d'évaluation est qu'il y a un écart entre la première mise en œuvre et les subséquentes. Julien a constaté que la phase d'évaluation est plus difficile lors de la première mise en œuvre, du fait des imprévus et des réajustements en découlant. Sa façon de réagir à cette situation a été d'être plus flexible dans son évaluation lors de la première implantation, afin que les étudiantes et étudiants ne subissent pas les contrecoups de ces ajustements. Il mentionne toutefois que la qualité et la rigueur de l'évaluation s'améliorent au fil des expériences de mise en

œuvre d'une même ÉTCSA, et il suggère de répéter souvent la même ÉTCSA afin de permettre une amélioration des outils d'évaluation et de permettre à l'enseignante ou l'enseignant d'avoir des points de repère stables.

Le dernier enjeu associé à la phase d'évaluation mentionné par Julien concerne les le temps et les ressources. Selon son expérience, outre la première mise en œuvre, la phase d'évaluation depuis son adoption de l'ÉTCSA ne demande pas plus de temps que lors d'une évaluation traditionnelle, mais elle est répartie sur une plus longue portion de la session. Ainsi, il suggère de répéter souvent le même concept d'ÉTCSA pour permettre d'amortir l'investissement nécessaire lors de la première implantation, et ainsi ne pas provoquer un alourdissement de la tâche pour cette phase de l'évaluation.

L'interprétation de ces résultats en regard de la littérature sur l'évaluation en situation authentique permet de constater que plusieurs des enjeux rencontrés par Julien avaient été identifiés dans les écrits. Ceux-ci ont été regroupés en quatre catégories : la qualité et la rigueur du processus évaluatif, les exigences en temps et en ressources, le contexte et les contraintes institutionnelles et les problèmes de nature pédagogique.

Sur le plan de la qualité et de la rigueur du processus évaluatif, les enjeux de fiabilité du fait de la grande hétérogénéité du produit de l'ÉSA sont soulevés à la fois par Julien et par plusieurs autrices et auteurs ayant étudié d'ÉSA. Toutefois, si la cause identifiée est similaire, la teneur des préoccupations est différente. Alors que dans la littérature, la question de la fiabilité est abordée principalement dans la perspective de la concordance des jugements entre les évaluateurs et de l'écart de jugement possibles entre différents évaluateurs. Conséquemment, les solutions proposées par ceux-ci concernaient les moyens d'atteindre une plus grande homogénéité entre les jugements portés sur le produit d'une ÉSA. Or, dans le cas des données recueillies, les principaux questionnements concernent davantage la fiabilité des évaluations selon les différents rôles et tâches. Ces préoccupations sont davantage abordées dans la littérature sur

l'évaluation des performances en générale, et non de la littérature spécifique à l'évaluation en situation authentique. Toutefois, chez Julien comme dans la littérature, la conclusion est la même devant ces problèmes de fiabilité, soit que le sacrifice d'une certaine part de fiabilité est acceptable au profit de la validité des évaluations, qui est supérieure aux évaluations plus traditionnelles.

Dans la littérature, la fiabilité des évaluations est aussi abordée à travers la question des biais induits par le contexte de réalisation et des problèmes de généralisation de l'évaluation à partir d'une seule situation de performance, dont on peut difficilement garantir qu'il est assurément représentatif de l'apprentissage global réalisé, qui sont questionnés. On retrouve également cette préoccupation dans les données recueillies, mais elle est formulée non pas en tant qu'enjeu de fiabilité, mais en tant qu'enjeu d'équité entre les étudiantes et les étudiants, notamment en raison du contexte de comparaison à travers des outils comme la Cote R. Ainsi, s'il y a convergence quant aux constats entre nos données et la littérature, soit la préoccupation découlant du caractère contextualisé de la tâche et de l'hétérogénéité des situations et des productions, cela n'est pas associé à la même dimension de l'évaluation (problème de fiabilité versus problème d'équité).

Finalement, une autre différence entre la littérature et les propos recueillis dans la présente recherche est que, dans la littérature, la question de la fiabilité est principalement abordée en lien avec la phase d'évaluation, alors que dans les données recueillies, elle concerne aussi la phase d'élaboration.

La question des exigences en temps et en ressources est présente dans la littérature, tout comme dans le matériel recueilli dans la présente recherche, et il y a, dans l'ensemble, concordance entre les éléments. La très large majorité des autrices et auteurs s'accorde sur le fait que la mise en place d'une ÉSA demande plus de ressources, et que ce besoin en ressources se fait principalement sentir lors de la phase d'élaboration de l'ÉSA. Les données recueillies dans la présente étude concordent

parfaitement avec ces éléments de la littérature: en effet, le fait de disposer d'une libération (dégrèvement) des tâches d'enseignement pour pouvoir concevoir son projet d'ÉTCSA comme le facteur le plus important dans la capacité de l'enseignante ou de l'enseignant à mener à bien la phase d'élaboration. Un autre élément de consensus entre la littérature et les données recueillies est que les ressources nécessaires à la conception d'un canevas d'ÉSA sont relativement incompressibles. Sur le plan des ressources associées à la phase d'administration, il y a concordance entre les constats se dégageant de la littérature et ceux relevant des données recueillies sur le fait que l'administration d'une ÉSA demande plus de soutien et d'accompagnement pédagogique, en raison de rétroaction fréquente nécessaire et de soutien accru aux étudiantes et étudiants.

Un élément présent dans les données recueillies qui n'a pas été identifié dans la littérature est qu'il y a une différence importante entre la première implantation, beaucoup plus exigeante, et les mises en œuvre subséquentes. Or, cette distinction n'est pas présente dans la littérature consultée. Autre écart entre la littérature et les données de la présente recherche: dans la littérature, le soutien nécessaire lors de la phase d'administration implique d'ajouter des ressources professionnelles ou de limiter le nombre d'étudiantes et étudiants par groupe. Outre le cas de la première mise en œuvre, les données recueillies ne sont pas aussi tranchées - la variation du nombre d'étudiantes et étudiants par classe n'ayant pas été identifiée comme ayant plus d'impact sur sa capacité à mener à bien la phase d'administration de son ÉTCSA que dans une approche traditionnelle de l'évaluation. Finalement, au niveau des ressources associées à la phase d'évaluation il y a un écart entre la littérature et l'expérience du participant : alors que pour plusieurs autrices et auteurs, la phase d'évaluation nécessite substantiellement plus de temps dans une approche en situation authentique que dans une approche psychoéducative de l'évaluation, les données recueillies indiquent plutôt que, outre la première expérience de mise en œuvre d'une nouvelle ÉTCSA, le temps nécessaire à la réalisation de la phase d'évaluation des productions des étudiantes et étudiants n'était pas supérieur à l'approche utilisée antérieurement. Toutefois, dans le cas de la présente étude, le point de référence n'était pas une approche

psychoéducative, mais plutôt une approche d'évaluation des performances, ce qui peut expliquer en partie cet écart.

Sur le plan du contexte et des contraintes institutionnelles, les quelques autrices et auteurs qui abordent ces enjeux traitent principalement des difficultés associées aux courants de standardisation des modalités d'évaluation, ayant encouragé la mise en place de mécanismes dont les normes sont parfois si précises qu'elles n'offrent pas la latitude nécessaire à la conception d'ÉSA, particulièrement lorsqu'elles impliquent une part importante de travail collaboratif. Sur ce point, les données recueillies ne démontrent pas d'entraves majeures relevant de cette tendance : les démarches qui ont été rapportées se sont bien déroulées et étaient relativement simples et rapides. Toutefois, la présence d'une ouverture au sein de l'institution face à une certaine variabilité des tâches pour une même évaluation a été identifiée par Julien comme centrale dans la possibilité de mettre sur pied un projet d'ÉTCSA. Ainsi, les conditions identifiées comme favorables dans la littérature étaient réunies dans le cas à l'étude dans la présente recherche, et cela a bel et bien été une condition favorable identifiée dans les données recueillies.

Dans la présente recherche, deux éléments relevant du contexte institutionnel non documenté dans la littérature ont été recensés, relevant spécifiquement du contexte institutionnel des études collégiales : 1) le format des cours, qui restreignait les possibilités quant à l'authenticité du contexte de réalisation, et 2) le système d'admission universitaire basé sur la cote R, qui place les étudiantes et étudiants en compétition entre eux, et rend délicat le fait d'évaluer une même compétence de manière non standardisée au sein d'un même groupe. Ces découvertes étaient inattendues et amènent des éléments nouveaux pour une compréhension accrue des facteurs favorisant ou entravant l'implantation d'ÉTCSA.

Finalement, sur le plan des problèmes de nature didactique et pédagogique, les données recueillies rejoignent les deux grandes catégories d'enjeux abordés dans la

littérature : soit les accros à l'authenticité de la tâche conçue et les impacts du manque d'expérience des apprenants en matière d'ÉSA.

On retrouve dans la littérature une préoccupation concernant le degré d'authenticité de certaines tâches présentées comme des ÉSA, mais qui n'en aurait en fait que l'apparence en ajoutant une contextualisation sommaire et non les caractéristiques profondes. On peut ici établir un parallèle entre le choix du participant à notre étude de se rabattre sur une plateforme numérique pour la réalisation d'une portion de la tâche, du fait que le contexte institutionnel n'offre pas suffisamment d'heures en classe pour procéder dans un contexte le plus authentique possible. Ici, un enjeu relevant en premier lieu du contexte institutionnel engendre aussi un problème de nature pédagogique concernant le degré d'authenticité du contexte de réalisation de la tâche. Et bien que la cause soit différente dans l'étude de cas de celles identifiées dans la littérature (qui relève plutôt d'erreur de conception ou de compréhension du concept, ce qui n'est pas le cas chez le participant à cette étude, qui a opté consciemment pour ce scénario altérant l'authenticité de la tâche du fait que le scénario idéal était impossible dans le cadre qui lui était imparti), le résultat est le même, soit une altération du degré d'authenticité global de l'ÉTCSA. Et par le fait même, l'introduction d'un biais de validité dans la tâche et l'évaluation en résultant, tel qu'identifié par la littérature.

Finalement, en ce qui concerne les impacts du manque d'expérience des étudiantes et étudiants en matière d'ÉSA, faute d'avoir été souvent exposés à cette formule durant leur parcours scolaire, on retrouve bel et bien des traces dans les données recueillies de cet enjeu abordé dans la littérature. Et tout comme les autrices et auteurs consultés à ce propos, les données recueillies démontrent que cela engendre de l'insécurité et un grand besoin d'accompagnement chez les étudiants et étudiantes. Il ressort clairement des données qu'il y a un grand besoin d'accompagnement de la part des étudiantes et étudiants, pour des motifs similaires à ceux mentionnés dans la

littérature, soit la grande latitude dont disposent les étudiantes et étudiants, le caractère appliqué et inédit de la tâche et le niveau d'habileté intellectuel.

2. LES LIMITES DE LA DÉMARCHE

Cet essai présente des limites méthodologiques qu'il importe de souligner. Tout d'abord, sur le plan méthodologique, le choix d'une étude de cas unique, lequel a été sélectionné par un échantillonnage par choix raisonné, pose la limite de la non-représentativité. En effet, il n'est pas possible de généraliser le vécu du participant à cette étude à l'ensemble de ses collègues qui enseignent en sciences humaines. Toutefois, le choix de l'approche d'analyse des données, de nature qualitative, a été fait afin de traiter d'une manière appropriée pour ce type de données.

Autre préoccupation : la posture de la rédactrice de cet essai, qui possède elle-même les caractéristiques de la population à l'étude de par sa situation professionnelle. Cette proximité professionnelle peut engendrer des biais de confirmation ou des biais affectifs. La précaution de choisir un participant qui n'a pas de lien personnel ou professionnel immédiat a été privilégiée pour tenter de réduire ces impacts. Également, la triangulation des méthodes de collecte a été privilégiée afin de favoriser une posture la plus objective possible.

Finalement, mentionnons les limites inhérentes au choix de la principale méthode de collecte de données : l'entrevue semi-dirigée. En effet, malgré toutes les précautions prises lors de la rédaction du schéma d'entrevue et la réalisation de l'entrevue, il y a toujours une possibilité d'autocensure associée au biais de désirabilité sociale, présent avec ce type de méthode de collecte. Le choix de trianguler les méthodes de collecte, en ajoutant un questionnaire et en recueillant un dossier d'artefact, a contribué à pallier cette difficulté méthodologique.

3. LES PISTES POUR DES DÉMARCHES FUTURES

À la suite de la réalisation de cet essai, deux pistes d'approfondissement concernant les facteurs favorisant ou défavorisant la mise en œuvre d'évaluation en situation authentique nous semblent particulièrement intéressantes. La première serait la réalisation d'une recherche à partir de cas d'étude multiples en sciences humaines, qui permettrait de passer d'une perspective plus exploratoire à une visée de recherche descriptive plus complète. Ainsi, les présents résultats pourraient être réinvestis afin d'affiner le cadre théorique et produire des outils de collecte de données plus précis et plus performants. Il serait intéressant de continuer à cibler spécifiquement le programme de sciences humaines en raison de la place et de l'ampleur qu'occupe ce programme au sein du réseau collégial et de ses enjeux en termes de persévérance scolaire et de réussite, présentés précédemment.

Une seconde piste de recherche serait de s'intéresser à l'évaluation terminale de cours en situation authentique en décroisonner la population à l'étude, afin d'effectuer une analyse comparée entre plusieurs domaines d'étude. Cette approche serait complémentaire à la première proposition, car elle permettrait d'enrichir la compréhension des facteurs favorables et défavorables à la mise en place d'une telle forme d'évaluation en décontextualisant le contexte disciplinaire d'enseignement.

Une troisième piste de recherche concerne la découverte de l'écart entre la littérature et les propos recueillis pour ce qui touche aux enjeux de fiabilité de l'évaluation. Il serait pertinent de mieux comprendre pourquoi les écrits recensés s'intéressent à cet enjeu en lien avec la phase d'évaluation, alors que dans les données recueillies dans la présente étude situent ces questionnements davantage dans la phase d'élaboration.

Ces trois pistes de recherche futures illustrent bien la contribution du présent essai, qui a permis une exploration du vécu en matière de mise en œuvre d'une

évaluation terminale de cours en situation authentique, et a permis de mettre en lumière des facteurs favorables et défavorables à travers le vécu d'un enseignant en sciences humaines. En plus de documenter certains éléments dans un contexte québécois, éléments pour lesquels il existe à ce jour encore peu de littérature, cela a permis de dresser un portrait qui pourra servir de tremplin à des recherches plus approfondies sur le sujet des ÉTCSA. Le développement de connaissances en évaluation des apprentissages dans le contexte des études postsecondaires est en plein essor : les prochaines années seront cruciales pour assurer également la percolation de ces connaissances auprès des praticiennes et praticiens de cette ordre d'enseignement.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Archbald, D. (1991). Authentic Assessment: Principles, Practices and Issues. *School Psychology Quarterly*, 6(4), 279–293.
- Archbald, D. et Newmann, F. (1988). *Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School*. Reston, VA: National Association of Secondary School Principals.
- Ashford-Rowe, K., Herrington, J., et Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, 39(2), 205–222.
- Bélanger, D.-C. et Tremblay, K. (2012). *Portrait actualisé des croyances et des pratiques en évaluation des apprentissages au collégial*. Rapport de recherche, Montréal: Collège de Maisonneuve.
- Boisvert, J. (2007). Les représentations sociales du collège et de la réussite chez les garçons en Sciences humaines au collégial. *Pédagogie Collégiale*, 20(3), 15–22.
- Boisvert, J. et Paradis, J. (2008). *La motivation chez les garçons et les filles en sciences humaines au collégial*. Rapport de recherche, Saint-Jean-sur-Richelieu: Cégep Saint-Jean-sur-Richelieu.
- Boisvert, M., Lacoursière, M., et Lallier, A. (2006). L'aventure collective du Renouveau de 1992 à 2004: une maturation engageante. In L. Héon, D. Savard et T. Hamel (dir.), *Les Cégeps: une grande aventure collective* (pp. 71–92). Québec: Presses de l'Université Laval.
- Bond, L., Moss, P., et Carr, P. (1996). Fairness in Large-Scale Performance. In G. Philips (dir.), *Technical issues in large-scale performance assessment* (pp. 117–134). Washington D.C: National Center for Education Statistics.
- Boutin, G. (2006). De la réforme de l'éducation au «renouveau pédagogique»: un parcours chaotique et inquiétant, *Argument*, 9(1).

- Brennan, R. L. (1996). Generalizability of Performance Assessment. In G. Philips (dir.), *Technical issues in large-scale performance assessment* (pp. 19–58). Washington D.C: National Center for Education Statistics.
- Carpentier, R., et Conseil supérieur de l'éducation du Québec. (2000). *La formation du personnel enseignant du collégial: un projet collectif enraciné dans le milieu*. Avis au ministre de l'éducation. Québec: Conseil Supérieur de l'Éducation. Document téléaccessible à l'adresse <http://www.cse.gouv.qc.ca/fichiers/documents/publications/form_col.pdf>
- Chbat, J. (2004). *Les attitudes et les pratiques pédagogiques du collégial*. Rapport de recherche, Montréal: Collège André-Grasset.
- Chbat, J. (2005). Les attitudes et les pratiques pédagogiques des enseignants au collégial. *Pédagogie Collégiale*, 19(1), 15–22.
- Chbat, J. (2006). Les attitudes et les pratiques pédagogiques des enseignants du collégial. In L. Sylvain (dir.), *Recherches sur les pratiques pédagogiques au collégial* (pp. 23–30). Montréal: Chenelière éducation.
- Clarke, M. M., Madaus, G. F., Horn, C. L., et Ramos, M. A. (2000). Retrospective on educational testing and assessment in the 20th century. *Journal of Curriculum Studies*, 32(2), 159–181.
- Collège Édouard-Montpetit. (2013). *Politique sur l'éthique de la recherche avec des êtres humains*. Longueuil: Collège Édouard-Montpetit.
- Cronin, J. F. (1993). Four Misconceptions about Authentic Learning. *Educational Leadership*, 50(7), 78–80.
- Cumming, J. J., & Maxwell, G. S. (1999). Contextualising Authentic Assessment. *Assessment in Education*, 6(2), 177–194.
- Custer, R. L. (2000). Authentic assessment-Basic definitions and perspectives. *Using Authentic Assessment in Vocational Education*, (Information series no. 381), 1–5.
- Doirion, M.-M., Mercier, J., Pineault, M.-C., et Pineau, M. (2004). L'évaluation en situation authentique. Rimouski: Cégep de Rimouski. Développement pédagogique
- Durand, M.-J., et Chouinard, R. (2012). *L'évaluation des apprentissages. De la planification de la démarche à la communication des résultats*. Montréal: Marcel Didier.

- Duval, A.-M., et Pagé, M. (2013). *La situation authentique: de la conception à l'évaluation*. Montréal: Association québécoise de pédagogie collégiale.
- Ewing, S. C. (1998). Alternative Assessment: Popularity, Pitfalls, and Potential. *Assessment Update*, 10(1), 11–12.
- Fédération des cégeps et Institut du Nouveau Monde. (2008). *Les cégeps et leur milieu: défi attente et besoin* (p. 29). Montréal: Fédération des cégeps.
- Fortier, C. (2003). Les yeux grand fermés. *Pédagogie Collégiale*, 17(2), 9–14.
- Fortin, M.-F. (2010). *Fondements et étapes du processus de recherche* (5e édition.). Montréal: Chenelière éducation.
- Fréchette, N., et Mignault, C. (2010). *Recherche sur la persévérance scolaire des étudiantes et étudiants du programme de sciences humaines*. Rapport de recherche. Longueuil: Collège Édouard-Montpetit.
- Frey, B. B., Schmitt, V. L., et Allen, J. P. (2012). Defining authentic classroom assessment. *Practical Assessment, Research, & Evaluation*, 17(2). 1-18
- Gagnon, Y.-C. (2012). *L'étude de cas comme méthode de recherche* (2e édition.). Québec: Presses de l'Université du Québec.
- Gauthier, B. (2010). La structure de la preuve. In B. Gauthier (dir.), *Recherche sociale: de la problématique à la collecte des données* (5e édition., pp. 169–198). Québec: Presses de l'Université du Québec (1re éd. 1984).
- Gerard, F.-M. (2008). Évaluation des compétences: guide pratique. Document téléaccessible à partir de l'adresse <<http://catalogue.cdc.qc.ca/CDC/Pages/Main/Detail/Page.fr.aspx?PrmNoSeq=19&TypeRef=Normale&Prov=Normale&Ret=Normale&NoBnq=3&NoRef=6>>
- Goulet, J. P. (1995). L'évaluation sommative des compétences: un "beau" problème. *Enseigner Au Collégial*, Montréal, Association Québécoise de Pédagogie Collégiale, 399–405.
- Goulet, J.-P. (1994). Pour en finir avec la réforme. *Pédagogie Collégiale*, 8(2), 18–25.
- Gouvernement du Québec. (1995). *Évaluation des apprentissages et renouveau de l'enseignement collégial*. Québec: Commission d'évaluation de l'enseignement collégial.

- Gouvernement du Québec. (2004). *L'exercice des responsabilités dans les collèges: une première évaluation institutionnelle*. Rapport synthèse. Québec: Commission d'évaluation de l'enseignement collégial.
- Gouvernement du Québec. (2008). *Évaluation de programmes du renouveau de l'enseignement collégial*. Québec: Commission d'évaluation de l'enseignement collégial.
- Gouvernement du Québec. (2012). *L'application des politiques d'évaluation des apprentissages. Pour des évaluations justes et équitables*, Rapport synthèse. Québec: Commission d'évaluation de l'enseignement collégial.
- Gouvernement du Québec. (2013). *Statistiques de l'éducation. Enseignement primaire, secondaire, collégial et universitaire. Éditions 2011*. Québec: Ministère de l'éducation, des loisirs et du sport; Ministère de l'enseignement supérieur, de la science et de la technologie.
- Gresham, F. M. (1991). Alternative Psychometrics for Authentic Assessment ? *School Psychology Quarterly*, 6(4), 305–309.
- Gruda, A. (1994, 20 octobre). Cégep: l'approche par incompétence. *La Presse*, p. B2. Montréal.
- Gulikers, J. T. M. (2006). *Authenticity is in the Eye of the Beholder. Beliefs and perceptions of authentic assessment and the influence on student learning*. Thèse de doctorat. Open Universiteit Nederland, Maastricht. Document téléaccessible à http://www.ou.nl/documents/Promoties-en-oraties/Promoties/Promoties2006/Thesis_Judith_Gulikers_v4.pdf
- Guy, H. (s.d.). La description d'une situation authentique. Document non publié, téléaccessible à http://www.usherbrooke.ca/moodle-cours/file.php/5492/documents_du_cours/etape5/1258_instrument_6f.pdf
- Guy, H., Deshaies, P., et Poirier, M. (2004). L'évaluation des apprentissages. *Fédération des cégeps*. Document téléaccessible à http://www.usherbrooke.ca/moodle-cours/file.php/5492/documents_du_cours/etape3/Guy_et_al._2004_Chapitre-3_p._24-26pdf.pdf
- Haertel, E. H. (1999). Performance Assessment And Education Reform. *Phi Delta Kappan*, 80(9), 662–666.
- Haertel, E. H., et Linn, R. L. (1996). Comparability. In G. Philips (dir.), *Technical issues in large-scale performance assessment* (pp. 59–78). Washington D.C:

- National Center for Education Statistics. Document téléaccessible à <http://nces.ed.gov/pubs/96802.pdf>
- Hardy, R. (1996). Performance Assessment: Examining the Costs. In M. B. Kane et R. Mitchell (dir.), *Implementing Performance Assessment: Promises, Problems and Challenges* (pp. 107–118). Mahwah, N. J.: Lawrence Erlbaum Associates.
- Herrington, J., et Herrington, A. (2006). Authentic conditions for authentic assessment: Aligning task and assessment. In *29th Higher Education Research and Development Society of Australasia (HERDSA) Annual Conference*, (p.146-151) Milperra, New South Wales, Document téléaccessible à <http://researchrepository.murdoch.edu.au/5755/>
- Herrington, J., et Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, 48(3), 23–48.
- Hoepfl, M. (2000). Large-Scale Authentic Assessment. *Using Authentic Assessment in Vocational Education*, (381), 49–67.
- Howe, R., et Ménard, L. (1993). *Croyances et pratiques en évaluation des apprentissages*. Rapport de recherche, Laval: Collège Montmorency. Document téléaccessible à <http://www.cdc.qc.ca/parea/702242-howemenard-croyances-pratiques-evaluation-montmorency-PAREA-1993.pdf>
- Isabel, B. (2000). *Les changements de pratiques d'évaluation des apprentissages chez les enseignants de philosophie et de français dans le contexte du renouveau de l'enseignement collégial: une étude de cas dans un collège*. Thèse de doctorat en éducation, Université du Québec à Rimouski, Rimouski.
- Jalbert, P., Munn, J., et Dodier, P. (2001). *Liste de vérification des caractéristiques d'une situation authentique*. S.I: Société GRICS. Document téléaccessible à <http://recitmst.qc.ca/SAO/listeverificationprojet.pdf>
- Kamphaus, R. (1991). Authentic Assessment and Content Validity. *School Psychology Quarterly*, 6(4), 300–304.
- Karsenti, T., et Savoie-Zajc, L. (2011). *La recherche en éducation: étapes et approches* (3e édition.). Saint-Laurent: ERPI (1re éd. 2000).
- Kerka, S. (1993). Techniques for Authentic Assessment. ERIC Clearhouse on Adult, Career, and Vocational Education.
- Khattari, N., Reeve, A. L., et Kane, M. B. (1998). *Principles and Practices of Performance Assessment*. Mahwah, N. J.: Lawrence Erlbaum Associates.

- Laberge, M. (2010). *Le Développement d'un outil d'évaluation des apprentissages dans le cours "L'entreprise dans le monde contemporain"*. Essai de maîtrise en éducation. Université de Sherbrooke, Sherbrooke. Document téléaccessible à <http://www.cdc.qc.ca/pdf/laberge-evaluation-apprentissages-essai-usherbrooke-2010.pdf>
- Langevin, L., Boily, M., et Talbot, N. (2004). Les impacts pédagogiques de la réforme: qu'en pensent les enseignants. *Pédagogie Collégiale*, 17(4), 33–38.
- Legendre, R. (2005). *Dictionnaire actuel de l'éducation* (3e édition.). Montréal: Gérin.
- Leroux, J. L. (2010). *Comment des enseignants du collégial évaluent-ils des compétences? Rapport de recherche*. Saint-Hyacinthe: Cégep de Saint-Hyacinthe. Document téléaccessible à <http://www.cdc.qc.ca/parea/787402-leroux-evaluation-competences-st-hyacinthe-article-PAREA-2010.pdf>
- Louis, R. (2004). *L'évaluation des apprentissages: théories et pratiques*. Laval: Beauchemin. Retrieved from http://www.usherbrooke.ca/moodle-cours/file.php/5492/documents_du_cours/etape5/LOUIS-chapitre_7.pdf
- Lussier, O. (2004). La pédagogie de la première session: vers une nouvelle perception du rôle de l'enseignant. *Pédagogie Collégiale*, 17(4), 23–27.
- Lussier, O., et Allaire, H. (2004). L'évaluation «authentique». *Pédagogie Collégiale*, 17(3), 30–31.
- Madaus, G. F., et O'Dwyer, L. M. (1999). A Short History Of Performance Assessment. *Phi Delta Kappan*, 80(9), 688–695.
- Maeroff, G. I. (1991). Assessing Alternative Assessment. *Phi Delta Kappan*, 43(4), 272–281.
- McMann, F. C., et Jepsen McMann, C. (1992). Authentic Evaluation in History. *OAH Magazine of History*, 6(4), 6–9.
- Messick, S. (1994). The Interplay of Evidence and Consequences in the Validation of Performance Assessments. *Educational Researcher*, 23(2), 13–23.
- Messick, S. (1996). Validity in Performance Assessment. In G. Philips (dir.), *Technical issues in large-scale performance assessment* (pp. 1–18). Washington D.C: National Center for Education Statistics. Document téléaccessible à <http://nces.ed.gov/pubs/96802.pdf>

- Meyer, C. (1992). What's the difference between authentic and performance assessment. *Educational Leadership*, 49(8), 39–40.
- Monk, D. H. (1996). Conceptualizing the Cost of Large-Scale Pupil Performance Assessment. In M. B. Kane et R. Mitchell (dir.), *Implementing Performance Assessment: Promises, Problems and Challenges* (p. 119–137). Mahwah, N. J.: Lawrence Erlbaum Associates.
- Montgomery, K. (2002). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching*, 50(1), 34–40.
- Newmann, F. (1997). Authentic Assessment in Social Studies: Standards and Examples. In G. D. Phye (dir.), *Handbook of Classroom Assessment: Learning, Achievement, and Adjustment* (p. 359–380). San Diego: Academic Press.
- Newmann, F., et Archbald, D. (1992). The Nature of Authentic Academic Achievement. In *Toward a New Science of Educational Testing and Assessment* (p. 71–83). Albany: State University of New York Press.
- Newmann, F., Brandt, R., et Wiggins, G. (1998). An Exchange of View on Semantics, Psychometrics, and Assessment Reform: A Close Look at “Authentic” Assessments”. *Educational Researcher*, 27(6), 19–22.
- Newmann, F., Marks, H. M., et Gamoran, A. (1996). Authentic Pedagogy and Student Performance. *American Journal of Education*, 104(4), 280–312.
- Newmann, F., et Wehlage, G. (1993). Five Standards of Authentic Instruction. *Educational Leadership*, 50(7), 8–12.
- Ouimet, M. (1994a). Confusion et improvisation. *La Presse*, 18 octobre. p. A1. Montréal.
- Ouimet, M. (1994b). Des cours écrits à toute vapeur par de petits comités. 19 octobre. *La Presse*, p. A1.
- Ouimet, M. (1994c). Une méthode vague et un peu nébuleuse. 20 octobre. *La Presse*, p. A1.
- Ouimet, M. (1994d). Lucienne Robillard défend les grandes lignes de son projet de réforme des cégeps. 21 octobre. *La Presse*, p. C10.
- Paillé, P. (2007). La méthodologie de recherche dans un contexte professionnalisant: douze devis exemplaires. *Recherches Qualitatives*, 27(2), 133–151.

- Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment Research & Evaluation*, 13(4), 1–11.
- Pineault, M.-C. (2001). *Pratiques pédagogiques et approche par compétences dans le programme Sciences humaines du cégep de Rimouski*. Mémoire de maîtrise. Université du Québec à Rimouski, Rimouski.
- Popham, J. (1993). Circumventing the High Costs of Authentic Assessment. *Phi Delta Kappan*, 74(6), 470–473.
- Prégent, R., Bernard, H., et Kozanitis, A. (2009). *Enseigner à l'université dans une approche-programme: guide à l'intention des nouveaux professeurs et chargés de cours*. Montréal: Presses internationales polytechnique. Document téléaccessible à http://www.usherbrooke.ca/moodle-cours/file.php/5492/documents_du_cours/etape5/Pregent_et_al._2009_p._137-_189_.pdf
- Riley, K., et Slater Stern, B. (1998). Using Authentic Assessment and Qualitative Methodology to Bridge Theory and Practice. *The Educational Forum*, 62, 178–185.
- Roy, S. N. (2010). L'étude de cas. In B. Gauthier (dir.), *Recherche sociale: de la problématique à la collecte des données* (5e édition., pp. 199–226). Québec: Presses de l'Université du Québec (1re éd. 1984).
- Scallon, G. (2004). L'évaluation des apprentissages dans une approche par compétences. Document téléaccessible à partir de l'adresse <<http://catalogue.cdc.qc.ca/CDC/Pages/Main/Detail/Page.fr.aspx?PrmNoSeq=19&TypeRef=Normale&Prov=Normale&Ret=Normale&NoBnq=3&NoRef=13>>
- Schell, J. W. (2000). Think about Authentic Learning and Then Authentic Assessment. *Using Authentic Assessment in Vocational Education*, (381), 7–18.
- Scheurman, G., et Newmann, F. (1998). Authentic Intellectual Work in Social Studies: Putting Performance before Pedagogy. *Social Education*, 62(1), 23–25.
- Stecher, B. (2010). Performance Assessment in an Era of Standards-based Educational Accountability. Stanford University, Stanford Center for Opportunity Policy in Education. Document téléaccessible à partir de l'adresse <<https://scale.stanford.edu/system/files/performance-assessment-era-standards-based-educational-accountability.pdf>>

- Tanner, D. E. (2001). Authentic Assessment: A Solution, or Part of the Problem. *The High School Journal*, 85(1), 24–29.
- Tardif, J. (1993). L'évaluation dans le paradigme constructiviste. In R. Hivon (dir.) *L'évaluation Des Apprentissages. Réflexions, Nouvelles Tendances et Formation*. Sherbrooke: Editions Du CRP, 27–56.
- Taylor, C. (1994). Assessment for Measurement or Standards: The Peril and Promise of Large-Scale Assessment Reform. *American Educational Research Journal*, 31(2), 231–262.
- Terwilliger, J. S. (1997). Semantics, Psychometrics, and Assessment Reform: A Close Look at “Authentic” Assessments. *Educational Researcher*, 26(8), 24–27.
- Tremblay, G. (1995). D'une réforme structurelle à un renouveau de l'enseignement collégial. *Pédagogie Collégiale*, 9(1), 28–32.
- Vanasse, P. (2010). La simulation d'une situation authentique en science politique. In *Actes du 30e colloque de l'AQPC* (p. 73–80). Sherbrooke: Association québécoise de pédagogie collégiale. Document téléaccessible à http://www.cdc.qc.ca/actes_aqpc/2010/313.pdf
- Wiggins, G. (1989a). A True Test: Toward More Authentic and Equitable Assessment. *The Phi Delta Kappan*, 70(9), 703–713.
- Wiggins, G. (1989b). Teaching to the (Authentic) Test. *Educational Research Review*, 45(7), 41–47.
- Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research & Evaluation*, 2(2). Document téléaccessible à <http://pareonline.net/getvn.asp?v=2&n=2>
- Wiggins, G. (1993a). *Assessing Student Performance. Exploring the Purpose and Limits of Testing*. San Francisco: Jossey-Bass.
- Wiggins, G. (1993b). Assessment: Authenticity, Context, and Validity. *The Phi Delta Kappan*, 75(3), 210–214.
- Wiggins, G. (1998). *Educative Assessment: Designing Assessment to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.
- Wiggins, G. (2011). A true test. *Phi Delta Kappan*, 70(9), 703–713.

ANNEXE A

FORMULAIRE D'INFORMATION ET DE CONSENTEMENT

Partie 1 : Information

Titre du projet de recherche :

Les conditions favorisant la mise en œuvre d'une évaluation en situation authentique en sciences humaines au collégial

Chercheur(s) responsable(s) du projet de recherche (coordonnées) :

Johanne Paquin, professeure de sc. politique au Cégep Édouard-Montpetit, 945, ch. de Chambly, Longueuil, J4H 3M6. Courriel : johanne.paquin@cegepmontpetit.ca.
Téléphone : 450-679-2631 poste 5767.

Sous la supervision de : Julie-Lyne Leroux, professeure adjointe, Département de Pédagogie, Secteur PERFORMA, Université de Sherbrooke, Campus de Longueuil, Faculté d'éducation, 150, place Charles-Le Moyne, Bureau 200, Longueuil (Québec), J4K 0A8, Tél. 450 463-1835, poste 61742, julie.lyne.leroux@usherbrooke.ca,

Préambule

Je sollicite votre participation à un projet de recherche. Cependant, avant d'accepter de participer à ce projet et de signer ce formulaire d'information et de consentement, veuillez prendre le temps de lire, de comprendre et de considérer attentivement les renseignements qui suivent.

Ce formulaire peut contenir des mots que vous ne comprenez pas. Nous vous invitons à poser toutes les questions que vous jugerez utiles au chercheur responsable du projet ou aux autres membres du personnel affecté au projet de recherche et à leur demander de vous expliquer tout mot ou renseignement qui n'est pas clair.

Nature et objectifs du projet de recherche

Cette recherche porte sur la mise en œuvre d'épreuves terminales de cours basées sur les principes de l'évaluation en situation authentique dans le cadre de cours de première année du programme d'études préuniversitaires en sciences humaines.

Notre recherche devrait recruter un participant ou une participante correspondant au profil suivant :

Le participant a expérimenté l'implantation d'au moins une évaluation terminale de cours conforme aux principes de l'évaluation en situation authentique :

- 1) élaborée et mise en œuvre de sur une base volontaire et individuelle
- 2) lors d'une session régulière qui n'est pas antérieure à 2010 et qui n'a pas été significativement perturbée par des absences de la part de l'enseignant (par exemple, un congé de maladie de plus de deux semaines) ou des événements majeurs (par exemple, des grèves importantes).
- 3) dans le cadre d'un cours du programme d'études préuniversitaires en sciences humaines prévu à la 1^{re} année du programme d'étude dans la grille de cours du cégep de provenance
- 4) dans un cégep public
- 5) dans des classes où le nombre d'étudiants inscrits au premier jour de classe était d'au moins 35 étudiants
- 6) dans le cadre de cours qui ne s'inscrivaient pas dans un projet pédagogique spécial pour lequel les étudiants avaient fait l'objet d'une sélection (par exemple, un cours où les étudiants sont sélectionnés pour la participation à un stage de mobilité à l'étranger)

Déroulement du projet de recherche

La participation requise de la part des personnes sollicitées implique :

- 1) La participation à une entrevue semi-dirigée d'environ 1h30, et possiblement à une seconde entrevue d'un maximum d'une heure, si la première entrevue n'a pas permis de couvrir tous les thèmes prévus. Ces entrevues seront enregistrées sur bande audio, uniquement pour l'usage de la chercheuse.
- 2) La remise de documents pédagogiques utilisés dans le cadre des expériences de mise en œuvre d'évaluations terminales basées sur les principes de l'évaluation en situation authentique dont il est question dans l'entrevue.
- 3) De répondre à un questionnaire d'environ 2 à 3 pages sur les expériences de mises en œuvre d'évaluations terminales de cours basées sur les principes de l'évaluation en situation authentique réalisées depuis 2010.

Risques associés au projet de recherche

Les principaux risques pour les participants et les participantes consistent à ceux habituellement reliés à la présentation publique de ses pratiques professionnelles, y compris d'éventuelles critiques. Toutefois, ce risque sera amoindri du fait que l'identité du participant ou de la participante et de l'institution à laquelle il ou elle appartient ne seront pas divulgués, directement ou indirectement.

Inconvénients associés au projet de recherche

Le principal inconvénient associé à la participation au projet de recherche est le temps qui devra être consacré à fournir les informations demandées et à participer aux entrevues. Voici un estimé du temps anticipé pour chaque élément : 1) rassembler les documents pédagogiques à remettre = entre 30 minutes et 1 heure;

2) répondre au questionnaire : environ 30 minutes; 3) participer à la première entrevue = environ 1h30 (maximum de 1h45) et, si nécessaire, participer à une seconde entrevue = maximum d'une heure.

La ou les entrevues seront réalisées au lieu de préférence du participant ou de la participante, afin de minimiser les déplacements de celui-ci ou de celle-ci.

Avantages

Il n'existe pas d'avantages directement reliés à la participation à la recherche. Néanmoins, les retombées possibles sont, sur le plan individuel, l'opportunité de réaliser un bilan de pratiques professionnelles exemplaires, et sur le plan collectif, de favoriser l'avancement des connaissances en matière de mise en œuvre d'évaluations en situation authentique par le partage de son expérience professionnelle. Aucune compensation n'est prévue en échange de la participation à cette recherche.

Participation volontaire et possibilité de retrait

Votre participation à ce projet de recherche est volontaire. Vous êtes donc libre de refuser d'y participer. Vous pouvez également vous retirer de ce projet à n'importe quel moment, sans avoir à donner de raisons, en faisant connaître votre décision au chercheur responsable du projet ou à l'un des membres du personnel affecté au projet. Si vous décidez de vous retirer de l'étude, les données qui proviendront de votre témoignage de même que les renseignements qui vous concernent seront détruits.

Votre consentement ne vous prive d'aucun droit au recours judiciaire en cas de préjudice lié aux travaux de recherche.

Confidentialité

Durant votre participation à ce projet, le chercheur responsable recueillera et consignera dans un dossier de recherche les renseignements vous concernant. Seuls les renseignements nécessaires pour répondre aux objectifs scientifiques de ce projet seront recueillis. Ces renseignements seront conservés sur des supports informatiques non reliés à un réseau (disque dur externe appartenant à la chercheuse et clé USB de sauvegarde, conservée sous clé entre les séances de travail), sécurisés à l'aide d'un mot de passe connu uniquement de la chercheuse, accessible au besoin par la professeure de l'Université de Sherbrooke, Mme Julie-Lyne Leroux, qui supervise cette démarche de recherche dans le cadre d'un essai réalisé en vue de l'obtention d'une Maîtrise en enseignement au collégial.

Tous les renseignements recueillis demeureront strictement confidentiels dans les limites prévues par la loi. Afin de préserver votre identité et la confidentialité des renseignements, vous ne serez identifié que par un numéro de code. La clé du code reliant votre nom à votre dossier de recherche sera conservée par le chercheur responsable.

Le chercheur responsable du projet utilisera les données à des fins de recherche dans le but de répondre aux objectifs scientifiques du projet décrits dans le formulaire d'information et de consentement ci-dessus. Ces données seront conservées pendant au moins 5 ans suivant le dépôt final de l'essai présentant les résultats de recherche, par le chercheur responsable.

Les données pourront être publiées dans des revues spécialisées ou faire l'objet de discussions scientifiques, mais il ne sera pas possible de vous identifier.

À des fins de protection, notamment afin de pouvoir communiquer avec vous rapidement, vos noms et prénoms, vos coordonnées et la date de début et de fin de votre participation au projet seront conservés pendant un an après la fin du projet dans un répertoire à part maintenu par le chercheur responsable.

Vous avez le droit de consulter votre dossier de recherche pour vérifier les renseignements recueillis, et les faire rectifier au besoin, et ce, aussi longtemps que le chercheur responsable du projet ou l'établissement détiennent ces informations. Cependant, afin de préserver l'intégrité scientifique du projet, vous pourriez n'avoir accès à certaines de ces informations qu'une fois votre participation terminée.

Identification des personnes-ressources

Si vous avez des questions concernant le projet de recherche ou si vous éprouvez un problème que vous croyez relié à votre participation au projet de recherche, vous pouvez communiquer avec le chercheur responsable du projet de recherche aux numéros suivants : 450-679-2631 poste 5767

Pour toute question concernant vos droits en tant que sujet participant à ce projet de recherche ou si vous avez des plaintes ou des commentaires à formuler vous vous pouvez communiquer avec le président du CER à l'adresse électronique suivante : [retirée pour des raisons de confidentialité, car elle permettait d'identifier l'institution de provenance du participant]».

Surveillance des aspects éthiques du projet de recherche

Même si les chercheurs ont tout mis en œuvre pour respecter les principes éthiques directeurs présentés par l'Énoncé de politique des trois conseils (gouvernement du Canada, 2010) ainsi que la [*Politique d'éthique de la recherche avec des êtres humains*](#) du CRL, vous pouvez communiquer avec la présidence du comité d'éthique et de la recherche du cégep pour signaler tout manquement à cet égard, faire valoir votre point de vue ou poser des questions:

Partie 2 : Consentement

J'ai pris connaissance du formulaire d'information et de consentement. Je reconnais qu'on m'a expliqué le projet, qu'on a répondu à mes questions et qu'on m'a laissé le temps voulu pour prendre une décision.

Je consens à participer à ce projet de recherche aux conditions qui y sont énoncées. Une copie signée et datée du présent formulaire d'information et de consentement m'a été remise.

Nom et signature du participant

ou de la participante à la recherche

Date

Je certifie qu'on a expliqué au participant à la recherche et, le cas échéant au tiers autorisé, les termes du présent formulaire d'information et de consentement, que l'on a répondu aux questions que le sujet de recherche avait à cet égard et qu'on lui a clairement indiqué qu'il demeure libre de mettre un terme à sa participation, et ce, sans préjudice.

Je m'engage, avec l'équipe de recherche, à respecter ce qui a été convenu au formulaire d'information et de consentement et à en remettre une copie signée au sujet de recherche.

Nom et signature du chercheur ou

de la chercheuse responsable du projet de recherche

Date

ANNEXE B

OUTIL DE COLLECTE : LE QUESTIONNAIRE

Présentation : Ce court questionnaire vise à mieux connaître votre parcours professionnel et vos expériences en matière d'évaluation en situation authentique. Il faut prévoir environ 15 minutes pour y répondre. Si vous avez des questions, n'hésitez pas à les poser. Une fois le questionnaire complété, nous débuterons l'entrevue.

Votre parcours professionnel :

1- Dans quelle discipline du programme de Sciences humaines enseignez-vous ?

2- Depuis combien d'années enseignez-vous au cégep dans une discipline des sc. humaines ? _____ ans

3- Au cours des 2 dernières années, quelle proportion de votre tâche d'enseignement a été réalisée auprès d'étudiants de première année du programme de sc. humaines ?

- ☐ Moins de 25%
- ☐ Entre 25% et 49 %
- ☐ Entre 50% et 74%
- ☐ 75% ou plus

4- Avez-vous déjà suivi des cours de pédagogie crédités?

- ☐ Oui (*passer à la question 5*)
- ☐ Non (*passer à la question 6*)

5- Si oui, dans quel contexte académique ?

- ☐ Dans le cadre d'un programme régulier (application ouverte à tous) d'une université québécoise
- ☐ Programme de formation continue ou offert via l'employeur (ex : Performa) en collaboration avec une université
- ☐ Dans le cadre de cours universitaires hors programmes
- ☐ Dans le cadre d'un bacc. ou d'une maîtrise disciplinaire intégrant un volet optionnel de pédagogie
- ☐ Autre : _____

6- Avez-vous déjà suivi des formations non créditées en pédagogie (journées pédagogiques, formations organisées par l'employeur ou par une organisation, colloque, etc.) ?

- ☐ Oui (*répondez aux questions 7 et 8*)
- ☐ Non (*passez à la question 9*)
- ☐ Ne sais pas

7- Si oui, dans quel(s) cadre(s) ? _____

8- Sur quel(s) thème(s) ? _____

Vos expériences d'implantation d'évaluations terminales de cours basées sur les principes de l'évaluation en situation authentique (ÉTCSA)

9- Dans les 2 dernières années, auprès de combien de groupes-classes avez-vous expérimenté des ÉTCSA (évaluations terminales de cours basées sur les principes de l'évaluation en situation authentique) dans le cadre de cours **de la première année du programme de Sciences humaines**?

- ☐ 1 à 2
- ☐ 3 à 5
- ☐ 6 à 9
- ☐ 10 et +

10- Dans combien de cours **de la première année du programme de Sciences humaines** avez-vous expérimenté des ÉTCSA ? *Par exemple, si vous avez réalisé une ÉTCSA dans deux groupes-classes suivant le cours 385-103, la réponse est «1 cours».*

- ☐ 1 cours
- ☐ 2 ou 3 cours
- ☐ 4 cours et plus

11- Comment estimez-vous votre niveau de connaissance à propos du concept d'«évaluation en situation authentique» ?

- ☐ Je n'ai aucune connaissance à propos de ce concept;
- ☐ J'ai un faible niveau de connaissance à propos de ce concept;
- ☐ J'ai un bon niveau de connaissance à propos de ce concept
- ☐ J'ai un excellent niveau de connaissance à propos de ce concept

12- Laquelle ou lesquelles de ces affirmations représente le mieux votre prise de contact avec le concept d'«évaluation en situation authentique» ? *Vous pouvez cocher plus d'une réponse.*

- ☐ Je n'ai jamais lu, étudié ou entendu parler de ce concept
- ☐ J'ai étudié ce concept dans le cadre d'un cours ou d'une formation en pédagogie (crédité ou non)
- ☐ J'ai lu de façon autonome à propos de ce concept
- ☐ J'ai entendu parler de ce concept par des collègues
- ☐ J'ai entendu parler de ce concept par un conseiller pédagogique

13- À partir de votre expérience d'enseignement des deux dernières années dans le cadre de cours de la première année du programme de Sciences humaines, à quel degré êtes-vous en accord ou non avec les énoncés suivants?

<i>La mise en place d'une évaluation terminale de cours en situation authentique :</i>		Fortement en désaccord	Assez en désaccord	Assez en accord	Fortement en accord	Ne sais pas
a) a eu un effet positif sur ma relation pédagogique avec les étudiants et étudiantes						
b) a eu un effet positif sur les résultats (notes) des étudiants et étudiantes						
c) a eu un effet positif sur la qualité des productions (travaux) des étudiants et étudiantes						
d) a eu un effet positif sur la qualité des apprentissages réalisés par les étudiants et étudiantes						
e) a eu un effet positif sur mon développement professionnel						
f) a occasionné une augmentation du nombre d'heures consacrées à la conception de l'épreuve terminale de cours (conception générale et production du matériel) :	i. lors de ma 1 ^{re} expérience d'implantation					
	iii. lors des expériences subséquentes					

g) a occasionné une augmentation du nombre d'heures consacrées à l'évaluation sommative et la notation des épreuves terminales de cours	i. lors de ma 1 ^{re} expérience d'implantation					
	ii. lors des expériences subséquentes					
h) a occasionné une augmentation du nombre d'heures consacrées au soutien des étudiants et étudiantes à l'extérieur des heures de cours						
i) a occasionné une augmentation de la tâche (explications, rétroactions, etc.) nécessaire à l'encadrement des étudiants et étudiantes à l'intérieur des heures de cours						
j) a occasionné une augmentation de la tâche (explications, rétroactions, etc.) nécessaire à l'encadrement des étudiants et étudiantes à l'extérieur des heures de cours						
k) a complexifié mon travail de professeurs à l'extérieur des heures de cours						
l) a complexifié mon travail de professeur durant les heures de cours						

14- Au cours des 2 dernières années, quel est votre degré de satisfaction globale vis-à-vis vos expériences d'évaluation terminale de cours basés sur les principes de l'évaluation en situation authentique réalisées auprès de classes de 1^{re} année du programme de Sciences humaines ?

- ☐ Fortement insatisfait
- ☐ Assez insatisfait
- ☐ Assez satisfait
- ☐ Fortement satisfait

Profil sociodémographique :

15- Êtes-vous un homme ou une femme ?

- ☐ Homme
- ☐ Femme

16- Quel est votre âge ? _____

17- Au cours des 2 dernières années, avez-vous travaillé à temps plein chaque session ?

☐ Oui

☐ Non

18- Sinon, quel était **en moyenne** votre % de tâche au cours des deux dernières années ?

☐ 0 à 25 %

☐ 26 à 50%

☐ 51 à 75 %

☐ + de 75 %

Je vous remercie pour d'avoir pris le temps de compléter ce questionnaire.

ANNEXE C

OUTIL DE COLLECTE : LE SCHÉMA D'ENTREVUE

Présentation de l'entrevue :

Avant de commencer l'entrevue :

Présentation : maintenant, nous allons débiter l'entrevue. Cette entrevue portera sur votre expérience d'implantation d'évaluation terminale de cours en situation authentique dans le cadre de cours de sc. humaines de la première année du programme. Elle a pour objectif de documenter les conditions favorisant la mise en œuvre réussie d'une épreuve terminale de cours s'appuyant sur les principes d'une évaluation en situation authentique dans des cours de première année du programme d'études préuniversitaires en Sciences humaines. La durée prévue est d'une heure trente et elle sera enregistrée et retranscrite afin de faciliter le travail d'analyse, mais que les bandes audios ne seront utilisées qu'aux fins de la présente recherche et seront détruites une fois l'essai évalué !

Rappel de la délimitation spatiotemporelle et contextuelle concernant l'objet de l'entrevue :

Dans le cadre de cette entrevue, lorsque vous répondrez aux questions, gardez en tête que vos réponses doivent concerner les expériences d'évaluation terminale de cours en situation authentique :

- 1) qui ont eu lieu dans le cadre de cours de 1^{re} année du programme de Sciences humaines;
- 2) conçues et mises en application de façon libre (volontaire) et individuelle;
- 3) qui ont eu lieu entre le 1^{er} janvier 2015 et le 31 décembre 2017;
- 4) dans des classes où le nombre d'étudiants et d'étudiantes inscrits au premier jour de classe était d'au moins 35 étudiants ou étudiantes;

- 5) Dans le cadre de cours qui ne s'inscrivaient pas dans un projet pédagogique spécial.

Avant de commencer les questions, je vous invite à consulter les définitions suivantes (présentées sur une feuille à part) :

1) *Évaluation terminale de cours* : On entend par évaluation terminale de cours une activité d'évaluation réalisée à la fin de la session, qui mesure l'intégration des savoirs, savoir-faire et savoir-être centraux du cours et permet ainsi d'attester, pour chaque étudiant, de la maîtrise des éléments de compétences du cours selon les standards visés.

Avez-vous des questions ou des besoins de précisions concernant ce qu'est une évaluation terminale de cours ?

2) *Évaluation en situation authentique* : l'évaluation en situation authentique est une forme d'évaluation des performances dans laquelle le caractère authentique de la tâche proposée et de son cadre de réalisation est accentué pour devenir une préoccupation de premier plan tout au long de la démarche évaluative

Avez-vous des questions ou des besoins de précisions concernant ce qu'est une évaluation terminale de cours ?

Nous sommes maintenant prêts à commencer l'entrevue

Déroulement de l'entrevue :

L'entrevue sera divisée en 3 parties. La première vise à connaître vos expériences en matière d'implantation d'évaluation terminale de cours en situation authentique. Ensuite, nous aborderons votre vécu en matière d'implantation d'une évaluation terminale de cours en situation authentique pour chacune des étapes de l'implantation, soit l'élaboration, l'administration et l'évaluation. Nous discuterons des conditions qui ont facilité ou rendu plus difficile vos tentatives d'implantation d'évaluation terminale de cours en situation authentique, des obstacles rencontrés et des conditions ayant facilité vos démarches. Finalement, nous aborderons certains thèmes présents dans la littérature concernant les obstacles et les pistes de solution propres à ce type d'évaluation.

Partie 1 – Expériences d’implantation d’évaluation terminale de cours en situation authentique

Nous allons maintenant aborder vos expériences d’implantation d’évaluation terminale de cours en situation authentique.

1- Tout d’abord, présentez-moi l’évaluation terminale de cours en situation authentique que vous avez fait réaliser le plus souvent dans le cadre d’un cours de première année du programme de Sciences humaines (s’il y a deux expériences réalisées le même nombre de fois, nous utiliserons celle réalisée le plus récemment).

Pour le reste de l’entrevue, nous référerons à cette évaluation terminale de cours en situation authentique par le terme “évaluation de référence”

2- Avez-vous fait d’autres mises en œuvre d’évaluations terminales de cours en situation authentique? Si oui, présentez-les.

Partie 2 – Vécu en matière d’implantation d’une évaluation terminale de cours en situation authentique lors de l’élaboration, l’administration et l’évaluation

Pour la seconde partie de l’entrevue, nous discuterons d’abord d’évaluation terminale de cours en situation authentique de référence, puis nous discuterons d’autres expériences d’évaluation terminale de cours en situation authentique s’il y a lieu.

Commençons par aborder l’étape de l’élaboration.

Cette étape désigne la phase de conception du projet d’évaluation en situation authentique et de rédaction du matériel pédagogique et des outils d’évaluation associés. Il s’agit donc de l’étape qui précède la présentation en classe et de la remise du matériel pédagogique et du matériel d’évaluation aux étudiants et étudiantes (voir le document *Présentation des définitions*, destiné au participant).

3- Lorsque vous réfléchissez à la toute première fois où vous avez élaboré votre évaluation terminale de cours en situation authentique de référence, quelles sont les difficultés que vous avez rencontrées ? Quels ont été les principaux obstacles au déroulement optimal de l’élaboration de cette première mouture de votre évaluation terminale de cours en situation authentique ? Comment avez-vous fait face à ces difficultés et obstacles rencontrés ?

4- Quels ont été les aspects qui se sont le mieux déroulés ?

5- Quelles ont été les principales conditions ayant facilité la première élaboration de votre évaluation terminale de cours en situation authentique?

6- Ces conditions favorables et ces obstacles rencontrés ont-ils persisté lors que vous avez travaillé à améliorer à cette première mouture de votre évaluation terminale de cours en situation authentique?

7- (Selon l'information fournie lors de la première partie de l'entrevue) Avez-vous rencontré le même type de difficultés lors de la conception d'autres expériences évaluation terminale de cours en situation authentique? Les conditions qui ont facilité la conception de ces autres évaluations terminales de cours en situation authentique étaient-elles similaires ?

Passons maintenant à la phase de l'administration.

Cette étape désigne la phase durant laquelle la tâche est présentée aux étudiants, puis réalisée par ceux-ci. Cela inclut toutes les phases de réalisation, y compris les moments de rétroaction formative, et se termine à la remise de la production finale ou au moment de réalisation de la performance, selon le type de tâche demandée (voir le document *Présentation des définitions*, destiné au participant).

8- Lorsque vous réfléchissez à la toute première fois où vous avez administré votre évaluation terminale de cours en situation authentique de référence, quelles sont les difficultés que vous avez rencontrées ? Quels ont été les principaux obstacles au déroulement optimal cette première administration ? Comment avez-vous fait face à ces difficultés et obstacles rencontrés ?

9- Quels ont été les aspects qui se sont le mieux déroulés ? Qu'est-ce qui a contribué au bon déroulement de cette première administration de votre évaluation terminale de cours en situation authentique?

10- Ces éléments ont-ils persisté lors que vous avez travaillé à améliorer à cette première expérience d'administration de votre évaluation terminale de cours en situation authentique?

11- (Selon l'information fournie lors de la première partie de l'entrevue) Avez-vous rencontré le même type de difficultés lors de l'administration d'autres évaluations

terminales de cours en situation authentique? Les conditions qui ont facilité l'administration de ces autres évaluations terminales de cours en situation authentique étaient-elles similaires ?

Finalement, abordons la phase de l'évaluation.

Cette étape désigne la phase durant laquelle l'évaluateur analyse l'ensemble de la production ou performance finale à partir de chaque indicateur afin de porter un jugement et sanctionner l'atteinte ou non des objectifs d'apprentissage et des standards visés et certifier de l'atteinte ou non de la/des compétence (s) associée (s) au cours (voir le document *Présentation des définitions*, destiné au participant).

12- Lorsque vous réfléchissez à la toute première fois où vous avez évalué les productions ou performances finales résultant de votre évaluation terminale de cours en situation authentique de référence, quelles sont les difficultés que vous avez rencontrées ? Quels ont été les principaux obstacles au déroulement optimal cette première évaluation ? Comment avez-vous fait face à ces difficultés et obstacles rencontrés ?

13- Quels ont été les aspects qui se sont le mieux déroulés ? Quels ont été les principaux éléments qui ont favorisé le bon déroulement de cette première évaluation des productions/performance résultant de votre évaluation terminale de cours en situation authentique de référence ?

14- Ces éléments ont-ils persisté lors que vous avez travaillé à améliorer à cette première expérience d'évaluation résultant de votre évaluation terminale de cours en situation authentique de référence ?

15- (Selon l'information fournie lors de la première partie de l'entrevue) Avez-vous rencontré le même type de difficultés lors de la phase de l'évaluation pour d'autres évaluations terminales de cours en situation authentique? Les conditions qui ont facilité l'évaluation de ces autres évaluations terminales de cours en situation authentique étaient-elles similaires ?

Partie 3 - Thèmes présents dans la littérature concernant les obstacles rencontrés et les conditions perçues comme favorables propres à l'évaluation en situation authentique

Thème 1 : Qualité et rigueur du processus évaluatif :

Nous allons maintenant parler des enjeux concernant la qualité et la rigueur de l'évaluation, c'est-à-dire des enjeux comme la fiabilité, la précision et l'équité de l'évaluation et la capacité réelle de celle-ci à mesurer le niveau de maîtrise recherché dans le cadre du cours.

16- Avez-vous rencontré des enjeux concernant la qualité et la rigueur de l'évaluation lors de vos expériences d'implantation d'une évaluation terminale de cours en situation authentique? Si oui, lesquels? Comment les avez-vous surmontés? Lesquels n'ont pas pu être surmontés? Sinon, y a-t-il des éléments qui vous ont aidé à prévenir les problèmes de qualité et de rigueur?

17- Y a-t-il eu une évolution en matière de qualité et de rigueur au fil de vos expériences de mise en œuvre d'évaluations terminale de cours en situation authentique?

Thème 2 : Le temps et les ressources

18- Lors de vos expériences d'implantation d'une évaluation terminale de cours en situation authentique, avez-vous rencontré des enjeux liés aux exigences en temps et en ressources? Si oui, lesquels? Comment les avez-vous surmontés? Lesquels n'ont pas pu être surmontés? Sinon, y a-t-il des éléments qui vous ont aidé à prévenir ou contourner d'éventuels problèmes liés aux exigences en temps et en ressources?

Si non abordé spontanément, relancer sur les éléments suivants (difficultés et pistes de solutions):

19- Avez-vous d'autres précisions à apporter concernant la charge de travail associée à la phase d'élaboration d'une évaluation terminale de cours en situation authentique

20- Avez-vous d'autres précisions à apporter concernant les besoins matériels et logistiques nécessaires à la mise en place d'un contexte authentique?

21- Avez-vous d'autres précisions à apporter concernant les exigences en temps et en ressource associés à la réalisation d'une évaluation terminale de cours en situation authentique?

22- Avez-vous d'autres précisions à apporter concernant l'ampleur de la rétroaction associée à la réalisation d'une évaluation terminale de cours en situation authentique?

Thème 3 : Le contexte et les contraintes institutionnelles

23- Lors de vos expériences d'implantation d'une évaluation terminale de cours en situation authentique, avez-vous rencontré des enjeux liés aux contraintes institutionnelles (par exemple, liés aux politiques du cégep ou du département) ? Si oui, lesquels ? Comment les avez-vous surmontés ? Lesquels n'ont pas pu être surmontés ? Sinon, y a-t-il des éléments qui vous ont aidé à prévenir ou contourner les problèmes liés aux contraintes institutionnelles ?

Thème 4 : Les problèmes de nature didactique et pédagogique

24- Lors de vos expériences d'implantation d'une évaluation terminale de cours en situation authentique, avez-vous rencontré des enjeux de nature didactique ou pédagogique ? Si oui, lesquels ? Comment les avez-vous surmontés ? Lesquels n'ont pas pu être surmontés ? Sinon, y a-t-il des éléments qui vous ont aidé à prévenir ou contourner les problèmes de nature didactique et pédagogique ?

Si non abordés spontanément, relancer sur les éléments suivants (difficultés rencontrées et conditions facilitantes):

25- Avez-vous rencontré des difficultés associées à la place que prend le contexte de l'évaluation terminale de cours en situation authentique versus les compétences à démontrer lors de l'implantation d'évaluations terminale de cours en situation authentique?

26- Avez-vous déjà rencontré des difficultés à concevoir des tâches dont le contexte est réellement authentique (et non simplement un artifice associé à une tâche non authentique) ? Avez-vous déjà eu l'impression que malgré vos efforts pour créer une tâche et un contexte authentique, certains aspects de l'évaluation terminale de cours n'atteignaient pas cet objectif ?

27- Avez-vous expérimenté des difficultés quant à la rédaction du matériel destiné aux étudiants et étudiantes quant à votre capacité de les guider adéquatement dans le processus d'une évaluation terminale de cours en situation authentique?

28- Comment vos étudiants ont-ils réagi à vos propositions d'évaluations terminales de cours en situation authentique? Quel était leur niveau de réceptivité ? Quels sont les commentaires positifs reçus de leur part ? Quels sont les commentaires négatifs reçus de leur part ?

Conclusion

29- Y a-t-il des éléments ou des enjeux, positifs ou négatifs, concernant l'implantation d'évaluations terminales de cours en situation authentique qui n'ont pas été abordés dont vous aimeriez parler ?

30- Quels sont les trois conseils que vous donneriez à un collègue qui souhaiterait implanter une évaluation terminale de cours en situation authentique dans le cadre d'un cours de première année en sc. humaines ?

Je vous remercie de votre participation.

Présentation des définitions

(Feuille destinée au participant)

1- *Évaluation terminale de cours* : On entend par évaluation terminale de cours une activité d'évaluation réalisée à la fin de la session, qui permet d'évaluer l'intégration des savoirs, savoir-faire et savoir-être centraux du cours et permet ainsi d'attester, pour chaque étudiant, de la maîtrise des éléments de compétences du cours selon les standards visés.

2- *Évaluation en situation authentique* : l'évaluation en situation authentique est une forme d'évaluation des performances dans laquelle le caractère authentique de la tâche proposée et de son cadre de réalisation est accentué pour devenir une préoccupation de premier plan tout au long de la démarche évaluative

3- *Phase de l'élaboration d'une évaluation*: Cette étape désigne la phase de conception du projet d'évaluation en situation authentique et de rédaction du matériel pédagogique et des outils d'évaluation associés. Il s'agit donc de l'étape qui précède la présentation en classe et de la remise du matériel pédagogique et du matériel d'évaluation aux étudiants.

4- *Phase de l'administration* : Cette étape désigne la phase durant laquelle la tâche est présentée aux étudiants, puis réalisée par ceux-ci. Cela inclut toutes les phases de réalisation, y compris les moments de rétroaction formative, et se termine à la remise de la production finale ou au moment de réalisation de la performance, selon le type de tâche demandée.

5- *Phase de l'évaluation* : Cette étape désigne la phase durant laquelle l'évaluateur analyse l'ensemble de la production ou performance finale à partir de chaque indicateur afin de porter un jugement et sanctionner l'atteinte ou non des objectifs d'apprentissage et des standards visés et certifier de l'atteinte ou non de la/des compétence (s) associée (s) au cours.

ANNEXE D

OUTIL DE COLLECTE : GRILLE DE COLLECTE DES ARTEFACTS

Note : Afin de garantir la confidentialité de l'identité du participant, certains titres ont été caviardés, car ils permettaient d'identifier la discipline enseignée par celui-ci

Code	Nom du document	Catégorie de matériel (ET, EN ou MX)	Type de document (C, D, É ou A)
A1	Questionnaire lié à la simulation [REDACTED]	MX	A
A2	Composition [REDACTED] : Groupe	MX	A
A3	Composition des groupes par nom	MX	C
A4	Votes [REDACTED]	MX	A
A5	Interventions par groupe	EN	É
A6	Autres [REDACTED] : Interventions [REDACTED]	ÉT	C
A7	Autres [REDACTED] : Préparation au comité [REDACTED]	ÉT	C
A8	Chef [REDACTED]	ÉT	C
A9	Chef [REDACTED] [REDACTED]	ÉT	C
A10	Chef [REDACTED]	ÉT	C
A11	[REDACTED] Bilan des comités [REDACTED]	ÉT	C
A12	[REDACTED] Bilan des comités [REDACTED]	ÉT	C
A13	[REDACTED]	ÉT	C
A14	[REDACTED] Discours d'ouverture	ÉT	C
A15	[REDACTED] Gestion du [REDACTED]	ÉT	C
A16	[REDACTED] Interventions [REDACTED]	ÉT	C
A17	Calendrier simulation	MX	A
A18	Plan de cours	MX	A

Légende : Type de matériel : Matériel destiné aux étudiants = ÉT; Matériel utilisé uniquement par l'enseignante ou l'enseignant = EN; Utilisé par les étudiants et l'enseignant=MX. Types de documents : Énoncé de travail ou consignes = C; Documents soutenant l'apprentissage (textes, notes de cours, documents présentant des notions, etc.) = D; Matériel d'évaluation= É; Autre = A;

ANNEXE E

TABLEAU 1 – SYNTHÈSE DES PROBLÈMES RELIÉS À LA CONCEPTION ET LA MISE EN ŒUVRE D’UNE ÉSA RECENSÉS DANS LA LITTÉRATURE

Tableau 1
Synthèse des problèmes reliés à la conception et la mise en œuvre d’une ÉSA recensés dans la littérature

Catégories de problème	Enjeux et difficultés recensées	Conditions et actions favorables documentées	Autrices et auteurs l’ayant documenté
Qualité et rigueur du processus évaluatif	Les problèmes de fiabilité : L’ÉSA pose des problèmes spécifiques en matière de fiabilité en raison de l’hétérogénéité du produit ou de la performance finale, situation nécessitant une formation, un entraînement et une concertation adéquate de la part des évaluateurs afin de limiter l’interférence des perceptions individuelles de ceux-ci.	<ul style="list-style-type: none"> - Former adéquatement les évaluatrices et évaluateurs; - Prévoir des périodes de concertation entre évaluateurs, notamment par la réalisation d’exercices de concordance des jugements; - En cas de manque de ressource permettant la mise en place de ces pistes de solution : accepter le sacrifice d’un haut niveau de fiabilité au profit d’un haut niveau de validité et des impacts positifs sur l’apprentissage. 	Bond, Moss, et Carr, 1996; Brennan, 1996; Haertel, 1999; Ewing, 1998; Gresham, 1991; Hoepfl, 2000; Kamphaus, 1991;

	Les problèmes de validité : le caractère généralisable de l'évaluation sur le plan certificatif en raison de la nature contextualisée de l'ÉSA : du fait du temps nécessaire à la réalisation d'une ÉSA, il n'est pas possible de multiplier ce type d'évaluation afin de s'assurer de l'étendu et de la transférabilité des connaissances et habiletés acquises.	Porter une attention particulière au choix des contextes et scénarios de base des ÉSA afin d'en assurer le caractère représentatif de l'apprentissage global réalisé par l'étudiante ou l'étudiant.	Maeroff, 1991; Messick, 1994, 1996; Montgomery, 2002; Riley et Slater Stern, 1998; Tanner, 2001; Taylor, 1994; Terwilliger, 1997, Wiggins, 1990, 2011;
	Les problèmes d'équité : le caractère contextualisé de l'ÉSA est propice à induire des enjeux d'équité quant à l'origine ethnique, socioéconomique et linguistique des étudiantes et des étudiants.	Porter une attention particulière au choix des contextes et scénarios afin de limiter les biais culturel ou linguistique, bien qu'il soit difficile de complètement les éliminer en raison du haut niveau de contextualisation que nécessite une ÉSA.	
Exigences en temps et en ressource	<ul style="list-style-type: none"> - La quantité de ressources nécessaires à la conception d'une ÉSA demeure élevée et peu compressible même si le nombre d'étudiantes et d'étudiants concerné est petit; - L'administration d'une ÉSA demande la mise en place d'un contexte authentique, ce qui peut impliquer l'utilisation de ressources matérielles et humaines supplémentaires, sans compter le temps; - Les ressources nécessaires pour les phases de l'administration et de l'évaluation sont récurrentes et incompressibles. 	<ul style="list-style-type: none"> - Augmenter les ressources humaines et financières associées à l'apprentissage et à l'évaluation; - Adopter une perspective globale des coûts et effets de l'introduction d'ÉSA dans l'éducation en les percevant avant tout comme un investissement. 	Ewing, 1998; Hardy, 1996; Hoepfl, 2000; Kerka, 1993; Khattri, Reeve et Kane, 1998; Maeroff, 1991; Monk, 1996; Montgomery, 2002; Tanner,

	L'analyse de la production ou performance finale d'une ÉSA demande beaucoup plus de temps et de ressources qu'un test standardisé.	- Augmenter les ressources humaines et financières associées à l'apprentissage et à l'évaluation; - Adopter une perspective globale des coûts et effets de l'introduction d'ÉSA dans l'éducation en les percevant avant tout comme un investissement.	2001; Taylor, 1994
	- L'administration d'une ÉSA demande un accompagnement pédagogique plus important - L'ampleur de la rétroaction exigée par le modèle de l'ÉSA implique d'importantes ressources durant et après le processus.	- Augmenter les ressources humaines et financières associées à l'apprentissage et à l'évaluation; - Adopter une perspective globale des coûts et effets de l'introduction d'ÉSA dans l'éducation en les percevant avant tout comme un investissement.	
Le contexte et les contraintes institutionnelles	Les politiques concernant les modalités d'évaluation de certaines institutions ainsi que certains mécanismes de contrôle et de standardisation de la qualité de l'enseignement et d'imputabilité n'offrent pas latitude nécessaire à l'implantation d'ÉSA.	Favoriser une approche centrée sur l'apport du contexte d'apprentissage plutôt que sur une perspective individuelle de l'apprentissage et de l'évaluation.	Clarke <i>et al.</i> , 2000; Herrington et Herrington, 2006
Les problèmes de nature pédagogique	La sous-représentation de certains éléments de la compétence : - Certains scénarios visant l'authenticité d'une situation peuvent engendrer un focus; disproportionné sur des éléments du contexte plutôt que sur les connaissances et habiletés à maîtriser.	- Conception soignée et réfléchie des scénarios d'apprentissage et d'évaluation visant l'atteinte des réelles caractéristiques et composantes d'une ÉSA; - L'ÉSA doit être composée de tâches et de critères d'évaluation pertinents en regard des connaissances reconnues dans le domaine d'expertise concerné;	Cumming et Maxwell, 1999; Herrington et Herrington, 2006; Newmann et Archbald,

	<ul style="list-style-type: none"> - Certains éléments de la réalité étudiée peuvent être sous-représentés dans la conception ou la réalisation de l'ÉSA; - L'erreur de pertinence : lorsque la tâche appelle des réponses non pertinentes pour mesurer les connaissances et habiletés visées par la compétence à évaluer. <p>Le camouflage : Certains scénarios visant l'authenticité d'une situation peuvent parfois être du camouflage, c'est-à-dire une fausse contextualisation visant à dissimuler le caractère non authentique d'une tâche.</p>	<ul style="list-style-type: none"> - L'ÉSA doit couvrir de façon équilibrée et représentative toutes les dimensions de la réalité étudiée. 	1992; Montgomery, 2000; Wiggins, 1993 <i>a</i> , 1993 <i>b</i>
	<p>Le manque d'expérience des étudiantes et des étudiants avec ce type d'évaluation peut entraîner de l'anxiété, un grand besoin d'accompagnement et des difficultés pour ceux-ci à performer adéquatement dans un tel contexte.</p>	<ul style="list-style-type: none"> - Exposer davantage les étudiantes et les étudiants à des ÉSA; - Augmenter les ressources en phase d'implantation. 	

ANNEXE F **TABLEAU 2 – RÉSUMÉ DES ENJEUX, DIFFICULTÉS, CONDITIONS ET ACTIONS PERÇUES COMME FAVORABLES PAR JULIEN POUR LA PHASE D’ADMINISTRATION DE L’ÉTCSA**

Tableau 2
Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d’administration de l’ÉTCSA

	Enjeux et difficultés rencontrées par Julien		Conditions et actions perçues comme favorables par Julien	
	Constats	Exemples	Constats	Exemples
Le besoin de temps	La nécessité de disposer de beaucoup de temps pour la phase d’élaboration est le principal enjeu identifié par Julien pour ce qui est de la phase d’élaboration.	Julien mentionne explicitement qu’il n’aurait pas été capable de concevoir son ÉTCSA s’il n’avait pas eu de temps réservé à cette tâche dans son temps de travail.	Disposer de libération (dégrèvement) afin de réaliser les tâches associées à l’élaboration d’une ÉSA est une condition favorable à l’élaboration de telles évaluations.	Julien a obtenu une libération de la part de son institution d’enseignement et d’un organisme pour élaborer son projet d’ÉTCSA.
L’interaction avec les	Travailler seul peut : - Entraîner des enjeux d’équité et d’applicabilité;	Julien a trouvé difficile d’évaluer seul l’équité entre les différentes tâches présentes dans son projet d’ÉTCSA.	La possibilité de discuter avec des collègues de la même discipline ou des collègues ayant	Julien a consulté des collègues de sa discipline, qui ont accepté de lui donner une rétroaction sur certaines dimensions de

	<ul style="list-style-type: none"> - Rendre plus difficile la recherche de solution à des problèmes rencontrés. 	<p>Il a aussi trouvé difficile d'évaluer seul la faisabilité d'une tâche, par exemple, en regard du temps nécessaire ou de la capacité des étudiantes et des étudiants à la réaliser.</p>	<p>expérimenté des ÉTCSA dans d'autres disciplines :</p> <ul style="list-style-type: none"> - Aide à valider ou améliorer la tâche et le matériel associé; - Aide à trouver des sources d'inspiration ou des solutions à des difficultés rencontrées. 	<p>son projet et lui ont fait certaines suggestions d'amélioration concernant le concept et le matériel de son ÉTCSA.</p> <p>Il a aussi participé à une journée pédagogique durant laquelle il a été inspiré par une expérience d'ÉTCSA présentée par des collègues d'une autre discipline, qui a en partie inspiré son propre projet.</p>
Disponibilité des ressources spécialisées	<p>Les tâches authentiques peuvent parfois impliquer des habiletés et connaissances multiples, qui ne sont pas toutes maîtrisées au même niveau par la professeure ou le professeur.</p> <p>Le contexte académique peut parfois nécessiter le recours à des outils, notamment technologiques, afin de contourner des problèmes</p>	<p>Julien a dû avoir recours à l'utilisation de plateformes informatiques pour contourner le problème de l'impossibilité de disposer du temps suffisant pour la réalisation de la tâche entièrement en classe.</p> <p>L'opérationnalisation de la tâche via la plateforme informatique adaptée demandait un niveau de maîtrise dont ne disposait pas Julien et a nécessité le</p>	<p>L'accès à des ressources spécialisées qui peuvent soutenir la professeure ou le professeur pour certaines dimensions logistiques est parfois essentiel au bon déroulement de l'ÉTCSA.</p>	<p>Julien a eu l'aide d'un aide pédagogique spécialisée en technologies de l'information et des communications pour le soutenir dans la conception et l'opérationnalisation du volet technologique de son ÉTCSA et dans la recherche de solutions aux problèmes rencontrés.</p>

	logistiques et rendre la tâche opérationnalisable dans le cadre d'un cours, lesquelles peuvent nécessiter le recours à des ressources spécialisées.	recours à un conseiller technopédagogique.		
Le contexte institutionnel et académique	<p>Le contexte institutionnel amène certaines contraintes qui peuvent parfois nécessiter des démarches pour permettre l'intégration d'ÉTCSA.</p> <p>Le contexte institutionnel et académique rend parfois difficile l'authenticité du contexte de la tâche.</p>	<p>Julien a dû faire modifier le plan-cadre du cours associé à son ÉTCSA pour le rendre compatible avec certaines dimensions de son ÉTCSA.</p> <p>Julien a diminué l'authenticité de certaines dimensions de la tâche (retrait de certains rôles) et du contexte de réalisation (utilisation d'une plateforme informatique) afin que la tâche soit réalisable dans le cadre d'un cours du collégial et des contraintes institutionnelles.</p>	<p>La réceptivité et le soutien de la part de l'institution face aux difficultés rencontrées et aux adaptations nécessaires sont déterminants pour une implantation réussie d'ÉTCSA.</p> <p>La professeure ou le professeur doit disposer de latitude pour expérimenter de nouveaux modèles.</p> <p>Il est parfois nécessaire de dévier du modèle théorique</p>	<p>Les intervenants dans le processus de modification du plan-cadre du cours ont été réceptifs à la démarche et aux demandes du participant, et la procédure a été complétée rapidement.</p> <p>L'institution n'a pas entravé l'initiative d'ÉTCSA de Julien, malgré le caractère novateur du projet et celui-ci a disposé d'une grande latitude lors de la conception et des réajustements de l'ÉTCSA.</p> <p>Julien a pris certaines libertés en regard du modèle théorique d'ÉSA</p>

			<p>d'ESA pour s'adapter au contexte institutionnel et académique.</p> <p>Les ressources disponibles au niveau institutionnel facilitent le travail d'élaboration de l'ÉTCSA.</p>	<p>afin de faciliter son implantation dans le contexte concret d'enseignement</p> <p>Julien a eu l'appui de l'institution pour une demande de subvention afin d'élaborer son projet d'ÉTCSA et a reçu le soutien d'un conseiller pédagogique pour les dimensions technologiques de son projet.</p>
--	--	--	--	--

ANNEXE G

TABLEAU 3 – RÉSUMÉ DES ENJEUX, DIFFICULTÉS, CONDITIONS ET ACTIONS PERÇUES COMME FAVORABLES PAR JULIEN POUR LA PHASE D'ADMINISTRATION DE L'ÉTCSA

<p style="text-align: center;">Tableau 3 Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'administration de l'ÉTCSA</p>				
	Enjeux et difficultés rencontrées par Julien		Conditions et actions perçues comme favorables par Julien	
	Constats	Exemples	Constats	Exemples
Les éléments propres à la première mise en œuvre				
Difficultés et besoins non anticipés des étudiantes et étudiants	Malgré une bonne préparation, il est impossible de tout prévoir et la survenue d'éléments non anticipés est inhérente à la première mise en œuvre d'une ÉTCSA, notamment concernant les difficultés et besoins de la part des étudiantes et étudiants.	Julien avait surestimé la maîtrise par les étudiantes et étudiants de certains outils technologiques, du fonctionnement général du cadre collégial et de l'approche évaluative proposée.	Dans la planification de la phase d'administration, il faut se réserver du temps pour réagir aux imprévus.	Julien a créé de nouveaux outils pour répondre aux besoins non anticipés des étudiantes et étudiants (par exemple, une grille de compilation des réponses des étudiantes et des étudiants au questionnaire initial

			<p>Disposer de ressources humaines supplémentaires en classe facilite la prise en charge des éléments non anticipés.</p>	<p>et des documents présentant des exemples de ce qui était demandé).</p> <p>Julien a réservé plus de temps en classe pour s'ajuster aux besoins des étudiantes et des étudiants.</p> <p>Lors de la première mise en œuvre de son ÉTCSA, Julien a pu bénéficier de l'aide d'un stagiaire durant les périodes de classe, ce qui a facilité le soutien aux étudiantes et aux étudiants.</p>
--	--	--	--	---

La maîtrise imparfaite de la démarche et du matériel de la part de l'enseignante ou l'enseignant	L'utilisation de nouveau matériel et d'une nouvelle approche de l'évaluation implique une maîtrise imparfaite de celles-ci par la professeure ou le professeur.	Malgré la planification minutieuse des outils et du déroulement, Julien n'avait pas anticipé la complexité de gestion de certains d'entre eux.	Adopter une attitude de flexibilité est nécessaire lors d'une première implantation : ne pas hésiter à faire des aménagements au besoin et accepter que lors de la première implantation, il peut y avoir des impacts sur la validité, la fidélité, voire l'équité, et prendre des mesures ponctuelles pour les atténuer dès la phase d'administration.	Lors de la première mise en œuvre de l'ÉTCSA conçue, Julien a revu certaines exigences et certaines dimensions de la tâche, afin de pallier certaines imperfections du matériel et certains imprévus dans le déroulement.
Les éléments récurrents ou contextuels				
Besoin accru de soutien et d'accompagnement de la part des étudiantes et étudiants	Étant donné la nature moins standardisée de l'ÉTCSA, l'administration de l'ÉTCSA implique plus de soutien et d'accompagnement individuel des étudiantes et étudiants, en classe et à l'extérieur de la classe.	Julien rapporte un plus grand nombre de questions et un besoin accru de validation de la part des étudiantes et des étudiants comparativement à l'approche traditionnelle et il doit consacrer plus de temps à y répondre, en classe et à l'extérieur de la classe.	Réserver du temps en classe pour la réalisation d'une partie des tâches associées à l'ÉTCSA. Prévoir du temps pour le soutien et les suivis à l'extérieur de la classe.	Julien a réservé plus de temps en classe pour la réalisation de tâches associées à l'ÉTCSA. Il a utilisé l'approche de la classe inversée pour voir certaines notions théoriques et libérer du temps en classe.

	Capacité réduite d'entraide entre les étudiantes et les étudiants étant donné la nature non standardisée des tâches et d'une certaine imprévisibilité dans le déroulement.	Julien a remarqué que le niveau d'entraide possible entre les étudiantes et étudiants est moins grand dans le cadre de l'ÉTCSA étant donné la moins grande standardisation des tâches à accomplir.	Adapter le matériel pédagogique.	Julien fait beaucoup de suivi à l'extérieur de la classe, notamment via la messagerie interne du cégep. Julien a élaboré des consignes très détaillées pour guider les étudiantes et étudiants dans les étapes de réalisation de leur tâche, et il a fourni des exemples pour illustrer concrètement celles-ci.
La variabilité du contexte pédagogique et les imprévus	La composition de la classe varie d'une session à l'autre et d'un groupe à l'autre.	Julien doit parfois adapter les rôles et les consignes en fonction de la variation du nombre et du profil des étudiantes et des étudiants.	Disposer d'une latitude individuelle pour faire face aux imprévus de la phase d'administration.	Julien a parfois modifié sur-le-champ des dimensions de la tâche ou des consignes, afin de s'adapter à une situation imprévue, telle l'abandon du

	La composition de la classe se modifie en cours de session.	Il y a des abandons de cours durant la session, ce qui perturbe l'organisation des rôles et des tâches au sein du groupe.	Démarrer les étapes de la tâche de l'ÉTCSA après quelques semaines de cours.	cours par une étudiante ou un étudiant ayant un rôle important dans la tâche. Julien a repoussé le démarrage des étapes de l'ÉTCSA de quelques semaines comparativement à sa planification initiale, moment où la classe est plus stable.
L'importance des ressources humaines et matérielles	Selon le projet d'ÉTCSA, des ressources matérielles peuvent être nécessaires à la mise en place du contexte de réalisation de la tâche.	Julien a eu besoin d'un accès à des laboratoires informatiques.	Une fois la première implantation réalisée, la phase d'administration ne nécessite pas beaucoup de ressources additionnelles.	Julien n'a pas ressenti le besoin de disposer de ressources additionnelles majeures pour la phase d'administration.

ANNEXE H

TABLEAU 4 – RÉSUMÉ DES ENJEUX, DIFFICULTÉS, CONDITIONS ET ACTIONS PERÇUES COMME FAVORABLES PAR JULIEN POUR LA PHASE D'ÉVALUATION DE L'ÉTCSA

<p style="text-align: center;">Tableau 4 Résumé des enjeux, difficultés, conditions et actions perçues comme favorables par Julien pour la phase d'évaluation de l'ÉTCSA</p>				
	Enjeux et difficultés rencontrées par Julien		Conditions et actions perçues comme favorables par Julien	
	Constats	Exemples	Constats	Exemples
Les enjeux d'équité	La non-standardisation des rôles entraîne des enjeux d'équité entre les étudiantes et les étudiants d'un même groupe lors de l'évaluation.	Julien utilise des grilles d'évaluation différentes selon le rôle et la tâche assumée par chacun.	Porter une attention particulière lors de la phase d'élaboration afin de prévenir les problèmes d'équité lors de la phase d'évaluation.	Julien a comparé systématiquement les rôles présents dans la tâche afin de s'assurer qu'ils étaient de même niveau
	Les comportements et les choix des étudiantes et des étudiants ont un impact sur les autres étudiantes et étudiants, et cela engendre des enjeux	Julien mentionne que la présence d'étudiantes ou d'étudiants plus dynamiques dans un groupe a un effet d'émulation qui		Julien a produit des grilles d'évaluation spécifiques à chaque rôle afin d'adapter l'évaluation aux spécificités de chaque rôle et ainsi évaluer avec précision.

	d'équité entre les groupes d'un même cours.	augmente la qualité globale des productions; à l'inverse, un groupe où les étudiantes et étudiants s'investisse moins dans l'activité favorise des productions de moins grande qualité.		
L'écart entre la première mise en œuvre et les subséquentes	<p>Lors de la première implantation, gérer l'impact des imprévus sur l'évaluation.</p> <p>La qualité et la rigueur de l'évaluation s'améliorent au fil des expériences de mise en œuvre d'une même ÉTCSA.</p>	<p>Julien mentionne avoir fait face à des difficultés lors de la phase d'évaluation de la première implantation en raison des impacts de certains imprévus ou d'éléments non anticipés sur la production finale.</p> <p>Julien mentionne qu'avec l'expérience, il peut clarifier les exigences et attentes, ce qui lui permet de mieux guider ses étudiantes et étudiants lors des mises en œuvre subséquentes de la phase d'administration et</p>	Répéter souvent la même ÉTCSA.	Julien constate que son aisance à évaluer les productions des étudiants augmente avec la répétition de la même ÉTCSA session après session.

		d'augmenter la qualité et la rigueur de l'évaluation.		
Les enjeux liés au temps et aux ressources	Outre la première mise en œuvre, la phase d'évaluation ne demande pas plus de temps que lors d'une évaluation traditionnelle, mais elle est répartie sur une plus longue portion de la session.	Alors que la phase d'évaluation était très concentrée dans le temps et très intense lorsque Julien avait une ÉTCSA de type «travaux de session», il consacre maintenant sensiblement le même nombre d'heures, mais celles-ci sont réparties sur une plus longue période.	Répéter souvent la même ÉTCSA.	Julien rapporte qu'après plusieurs mises en œuvre de la même ÉTCSA, le temps à consacrer à la phase d'évaluation était redevenu la même qu'en approche traditionnelle.

